CHAPTER 10

The Consumption Capital Asset Pricing Model

Chapter Outline

Intermediate Financial Theory.

10.1 Introduction

Our asset pricing models thus far have been either one-period models such as the CAPM or multiperiod but essentially static, such as the Arrow–Debreu model. In the latter case, even if a large number of future periods are assumed, all decisions, including security trades, take place at date zero. It is in this sense that the Arrow–Debreu model is static. Reality is different, however. Assets are traded every period, as new information becomes available, and decisions are made sequentially, one period at a time, all the while keeping in mind the fact that today's decisions influence tomorrow's opportunities and constraints. Our objective in this chapter is to capture these dynamic features and to price assets in such an environment.

Besides adding an important dimension of realism, another advantage of a dynamic setup is to make it possible to begin to draw the link between the financial markets and the real side of the economy. Again, strictly speaking, this can be accomplished within an Arrow–Debreu economy: firms' present and future state-contingent production and investment decisions could also be modeled as occurring at date zero. However, as we will see, the main issues require a richer dynamic context where real production and consumption decisions are not made once and for all at the beginning of time, but progressively, as time evolves.

The present chapter begins this process by placing the production side of the economy temporarily in the background and studying the asset pricing implications of the resulting equilibrium consumption and dividend series. Under the consumption capital asset pricing model perspective, it is the properties of an economy's equilibrium per capita consumption series that ultimately determine asset pricing relationships. There is nothing unusual here: in our earlier chapters it was the desire of investors to maximize their expected utility of consumption that led to the formation of asset demands and equilibrium asset prices, those prices then being uniquely identified with investors' equilibrium consumption. As a by-product of these endeavors, we will also revisit the notion of risk-neutral valuation, specializing it in a way that allows the theory to be judged by actual economic data.

10.2 The Representative Agent Hypothesis and its Notion of Equilibrium

10.2.1 An Infinitely Lived Representative Agent

To accomplish these goals in a model of complete generality (in other words, with many different agents and firms) where asset prices can be tractably computed is beyond the present capability of economic science. As an alternative, we will make life simpler by postulating many *identical* infinitely lived consumers. This allows us to examine the decisions of a *representative*, stand-in consumer and explore their implications for asset pricing. In particular, we will assume that the representative agent acts to maximize the expected present value of discounted utility of consumption over his entire, infinite lifetime:

$$\max E\left(\sum_{t=0}^{\infty} \delta^t U(\tilde{c}_t)\right)$$

where δ is his discount factor and $U(\ )$ his period utility function with $U_1(\ ) > 0$, $U_{11}(\ ) < 0$. This construct is the natural generalization to the case of infinite lifetimes of the preferences considered in our earlier two-period examples. Its use can be justified by the following considerations.

First, if we model the economy as ending at some terminal date T (as opposed to assuming an infinite horizon), then the agent's investment behavior will reflect this fact. In the last period of his life, in particular, the agent will stop saving, liquidate his portfolio, and consume its entire value. There is no real-world counterpart for this action as the real economy continues forever. Assuming an infinite horizon eliminates these terminal date complications. Second, it can be shown, under fairly general conditions, that an infinitely lived agent setup is formally equivalent to one in which agents live only a finite number of periods themselves, provided they derive utility from the well-being of their descendants (a bequest motive). This argument is detailed by Barro (1974).

Restrictive as it may seem, the identical agents' assumption can be justified by the fact that, in a competitive equilibrium with complete securities markets, there is an especially intuitive sense of a representative agent: one whose utility function is a weighted average of the utilities of the various agents in the economy. In Box 10.1, we detail the precise way in which one can construct such a representative individual and discuss some of the issues at stake.

10.2.2 On the Concept of a "No-Trade" Equilibrium

In a representative agent economy, we must, of necessity, use a somewhat specialized notion of equilibrium—a *no-trade equilibrium*. If, indeed, for a particular model specification, some security is in positive net supply, the equilibrium price will be the price at which the representative agent is willing to hold that amount—the total supply—of the security. In other specifications, we will price securities that do not appear explicitly—securities that are said to be in zero net supply. The prototype of the latter is an IOU type of contract: in a one-agent economy, the total net supply of IOUs must, of course, be zero.[1] In this case, if at some price the representative agent wants to supply (sell) the security, since there is no one to demand it, supply exceeds demand. Conversely, if at some price, the representative agent wants to buy the security (and thus no one wants to supply it), demand exceeds supply. Financial markets are thus in equilibrium, if and only if, at the

[1] An IOU (I owe you) is simply a promise to pay a specific amount of money on a specific date.

BOX 10.1 Constructing a Representative Agent

In order to illustrate the procedure for constructing a representative agent, let us return to the two-period ($t = 0, 1$) Arrow—Debreu economy considered in Chapter 9 and assume markets are complete. Without loss of generality, assume K agents and N states of nature at $t = 1$. Each agent k, $k = 1,2, \ldots, K$, solves:

$$\max U^k(c_0^k) + \delta^k \sum_{\theta=1}^{N} \pi_\theta U^k(c_\theta^k)$$
$$\text{s.t. } c_0^k + \sum_{\theta=1}^{N} q_\theta c_\theta^k \leq e_0^k + \sum_{\theta=1}^{N} q_\theta e_\theta^k$$

where the price of period 0 endowment is normalized to 1, and the endowments of a typical agent k are described by the vector $\begin{pmatrix} e_0^k \\ e_1^k \\ \vdots \\ e_N^k \end{pmatrix}$

Under very standard assumptions (cf. Chapters 1 and 9), the equilibrium allocation is Pareto optimal, and, at the prevailing Arrow—Debreu prices $\{q_\theta\}_{\theta=1,\ldots,N}$, supply equals demand in every market:

$$\sum_{k=1}^{K} c_0^k = \sum_{k=1}^{K} e_0^k, \text{ and}$$
$$\sum_{k=1}^{K} c_\theta^k = \sum_{k=1}^{K} e_\theta^k, \text{ for every state } \theta$$

Since this competitive equilibrium allocation is Pareto optimal, no one can be better off without making someone else worse off. One important implication of this property for our problem is that there exists some set of weights $(\lambda_1, \ldots, \lambda_K)$, which in general will depend on the distribution of initial endowments, such that the solution to the following problem gives an allocation that is identical to the equilibrium allocation:

$$\begin{aligned} \max \sum_{k=1}^{K} \lambda_k \left\{ U^k(c_0^k) + \delta^k \sum_{\theta=1}^{N} \pi_\theta U^k(c_\theta^k) \right\}&, \\ \text{s.t. } \sum_{k=1}^{K} c_0^k &= \sum_{k=1}^{K} e_0^k, \\ \sum_{k=1}^{K} c_\theta^k &= \sum_{k=1}^{K} e_\theta^k, \ \forall \theta \\ \sum_{k=1}^{K} \lambda_k &= 1; \ \lambda_k > 0, \ \forall k \end{aligned} \tag{10.1}$$

(*Continued*)

BOX 10.1 Constructing a Representative Agent (Continued)

Maximization problem (10.1) is interpreted to represent the problem of a benevolent central planner attempting to allocate the aggregate resources of the economy so as to maximize the weighted sum of the utilities of the individual agents. We proposed a similar problem in Chapter 9 in order to identify the conditions characterizing a Pareto optimal allocation of resources. Here we see problem (10.1) as suggestive of the form the representative agent's preference ordering, defined over aggregate consumption, can take (the representative agent is denoted by the superscript A):

$$\begin{aligned} U^A(c_0^A, c_\theta^A) &= U_0^A(c_0^A) + \sum_{\theta=1}^{N} \pi_\theta U^A(c_\theta^A), \text{ where} \\ U_0^A(c_0^A) &= \sum_{k=1}^{N} \lambda_k U_0^k(c_0^A) \text{ with } \sum_{k=1}^{K} c_0^k = \sum_{k=1}^{K} e_0^k \equiv c_0^A \\ U^A(c_\theta^A) &= \sum_{k=1}^{N} \lambda_k \delta_k U^k(c_\theta^A) \text{ with } \sum_{k=1}^{K} c_\theta^k = \sum_{k=1}^{K} e_\theta^k \equiv c_\theta^A, \text{for each state } \theta \end{aligned} \tag{10.2}$$

The above setup generalizes to as many periods as we like and, with certain minor modifications, to an infinite horizon. It accommodates future state-contingent endowments (e.g., representing uncertain labor income streams) as well as future production. Construct (10.2) presents an intuitive sense of a representative agent as one who constitutes a weighted average of all the economy's participants. Complete financial markets are a critical antecedent, as are expected utility preference representations. For more details, see Constantinides (1982), who first proposed this sense of a representative agent.

An important feature of construct (10.2) resides in the fact that, in general, the weights $\{\lambda_1, \lambda_2, \ldots, \lambda_K\}$ will depend on the initial endowments: Loosely speaking, the agent with more wealth gets a bigger λ weight. In effect, this feature of the λ weights means that different initial wealth distributions may give rise to "different representative agents." Nothing in this statement is surprising: different initial wealth distributions will give rise to different competitive equilibria, which must then be matched, via the procedure above to different representative agents.

Rubinstein (1974) shows, however, that the utility function, constructed in Eq. (10.2), will, in addition, be independent of the initial endowment distribution if three further conditions are satisfied:

1. The subjective time preference parameter of every agent is the same (i.e., all agents δ's are the same).
2. Agents' period preferences are identical and assume either of the two following forms:

$$U^k(c) = \frac{\gamma}{\gamma - 1}(\alpha^k + \gamma c)^{1-\frac{1}{\gamma}} \text{ or } U^k(c) = -e^{-\alpha^k c}$$

3. Agents receive no income or endowments in the period $t \geq 1$.

(*Continued*)

BOX 10.1 Constructing a Representative Agent (Continued)

If these conditions are met—i.e., by and large, if the agents' preferences can be represented by either a CRRA or a CARA utility function—then there exists a representative agent economy for which the equilibrium Arrow—Debreu security demands, and thus their equilibrium prices and rates of return, are the same as in the original heterogeneous agent economy.

In other words, the demand for the Arrow—Debreu security paying off in some arbitrary state $\hat{\theta}$ is of the form

$$z_{\hat{\theta}}\left(\{q_\theta\}_{\theta=1,2,\ldots,N}, \sum_{k=1}^{K} e_0^k\right)$$

In this event, "demand aggregation" is said to result.

What does a representative agent "look like" under Rubinstein (1974) aggregation? All agents in his economy have the same utility function with the same parameter, γ, the same $\alpha = (1/K)\sum_{k=1}^{K} \alpha^k$, the same δ, the same initial wealth $(1/K)\sum_{k=1}^{K} e_0^k$, and, accordingly, the same equilibrium consumption. See the Web Notes for Guvenen's (2011) summary statement of Rubinstein's (1974) aggregation result.

Constantinides's (1982) notion of a "representative agent" is more general than Rubinstein's (1974), principally in two ways: (1) there is no assumption that agent preferences are identical, or of some specific form, and (2) exogenous endowments or endogenous production in future date-states is admitted. The cost of this generality is a loss of the aggregation property: different initial endowment distributions lead to different associated "representative agents," and thus different equilibrium security prices and returns. In this chapter and in succeeding ones, our notion of a representative agent will be that of Constantinides (1982). Note that this choice demands both complete markets and VNM-expected utility preference representations.[35]

[35] All these comments notwithstanding, the simplest sense of a representative agent occurs when all agents are assumed to have (i) identical preferences, (ii) identical initial wealth, and (iii) identical future income shocks (which must be, by definition, aggregate shocks). They then construct the same portfolios and at all times undertake the same savings and possess the same wealth. Idiosyncratic income shocks are not allowed under this interpretation.

prevailing price, supply equals demand *and* both are simultaneously zero. In all cases, the equilibrium price is that price at which the representative agent wishes to hold exactly the amount of the security present in the economy. Therefore, the essential question being asked is: What prices must securities assume so that the amount the representative agent *must* hold (for all markets to clear) exactly equals what he *wants* to hold? At these prices, further trade is not utility enhancing. In a more conventional multiagent economy, an identical state of affairs is verified post-trade. The representative agent class of models is not appropriate, of course, for the analysis of some issues in finance. For example, issues linked with the volume of trade cannot be studied since, in a representative agent model, trading volume is, by construction, equal to zero.

10.3 An Exchange (Endowment) Economy

10.3.1 The Model

This economy will be directly analogous to the Arrow–Debreu exchange economies considered earlier: production decisions are in the background and abstracted away. It is, however, an economy that admits recursive trading, where investment decisions are made period by period (as opposed to being made once and for all at date 0).

There is one, perfectly divisible *share*, which we can think of as representing the market portfolio of the CAPM (later we shall relax this assumption). Ownership of this share entitles the owner to all the economy's output. (In this economy, all firms are publicly traded.) Output is viewed as arising exogenously and as being stochastically variable through time, although in a stationary fashion. This is the promised, though still remote, link with the real side of the economy. Indeed, we will use macroeconomic data to calibrate the model in the forthcoming sections. At this point, we can think of the output process as being governed by a large-number-of-states version of the three-state probability transition matrix found in Table 10.1.

That is, we assume there are a given number of output states, levels of output that can be achieved at any given date, and the probabilities of the transition from one output state to another are constant and represented by entries in the matrix **T**. The stationarity hypothesis embedded in this formulation may, at first sight, appear extraordinarily restrictive. The output levels defining the states may, however, be normalized variables, to allow for a constant rate of growth. Alternatively, the states could themselves be defined in terms of growth rates of output rather than output levels. See Appendix 10.1 for a growth illustration.

Table 10.1: Three-state probability transition matrix

Output in Period $t+1$

$$\text{Output in Period } t \qquad \begin{array}{c} \\ Y^1 \\ Y_2 \\ Y^3 \end{array} \begin{array}{c} \begin{array}{ccc} Y^1 & Y^2 & Y^3 \end{array} \\ \begin{bmatrix} \pi_{11} & \pi_{12} & \pi_{13} \\ \pi_{21} & \pi_{22} & \pi_{23} \\ \pi_{31} & \pi_{32} & \pi_{33} \end{bmatrix} \end{array} = T$$

where $\pi_{ij} = \text{Prob}(Y_{t+1} = Y^j; Y_t = Y^i)$ for any t.

In the continuous-state version of this perspective, the output process can be analogously described by a probability transition *function*

$$G(Y_{t+1}|Y_t) = \text{Prob}(Y_{t+1} \leq Y^j; Y_t = Y^i)$$

We imagine the security as representing ownership of a fruit tree where the (perishable) output (the quantity of fruit produced by the tree—the dividend) varies from year to year. This interpretation is often referred to as the *Lucas tree* economy in tribute to 1995 Nobel Prize winner, R.E. Lucas, Jr., who, in his 1978 article, first developed the consumption capital asset pricing model (CCAPM). The power of the approach, however, resides in the fact that any mechanism delivering a stochastic process on aggregate output, such as a full macroeconomic equilibrium model, can be grafted on the CCAPM, thus allowing an in-depth analysis of the rich relationships between the real and the financial sides of an economy.

Ours will be a *rational expectations economy*. By this expression, we mean that the representative agent's expectations will be on average correct and, in particular, will exhibit no systematic bias. In effect we assume, in line with a very large literature (and with most of what we have done implicitly so far), that the representative agent knows both the general structure of the economy and the exact output distribution as summarized by the matrix T. One possible justification is that this economy has been functioning for a long enough time to allow the agent to learn the probability process governing output and to understand the environment in which he operates. Accumulating such knowledge is clearly in his own interest if he wishes to maximize his expected utility.

The agent buys and sells securities (fractions of the single, perfectly divisible share) and consumes dividends. His security purchases solve:

$$\max_{\{z_{t+1}\}} E\left(\sum_{t=0}^{\infty} \delta^t U(\tilde{c}_t)\right)$$
$$\text{s.t. } c_t + q_t^e z_{t+1} + 1 \leq z_t Y_t + q_t^e z_t$$
$$z_t \leq 1, \ \forall t$$

where q_t^e is the period t real price of the security in terms of consumption (the price of consumption is 1) and z_t is the agent's beginning-of-period t holdings of the security. Holding a fraction z_t of the security entitles the agent to the corresponding fraction of the distributed dividend Y_t, which, in an exchange economy without investment, equals total available output. The expectations operator applies across all possible values of Y feasible at each date t with transition probabilities provided by the matrix **T**.

Let us assume that the representative agent's period utility function is strictly concave with $\lim_{c_t \to 0} U_1(c_t) = \infty$. Making this latter assumption ensures that it is never optimal for the agent to select a zero consumption level. It thus normally ensures an interior solution to the relevant maximization problem. The necessary and sufficient condition for the solution to this problem is then given by: For all t, z_{t+1} solves:

$$U_1(c_t)q_t^e = \delta E_t\{U_1(\tilde{c}_{t+1})(\tilde{q}_{t+1}^e + \tilde{Y}_{t+1})\} \tag{10.3}$$

where $c_t = (q_t^e z_t + z_t Y_t - q_t^e z_{t+1})$. Note that the expectations operator applies across possible output state levels; if we make explicit the functional dependence on the output state variables, Eq. (10.3) can be written (assuming Y^i is the current state):

$$U_1(c_t(Y^i))q_t^e(Y^i) = \delta \sum_j U_1(c_{t+1}(Y^j))(q_{t+1}^e(Y^j) + Y^j)\pi_{ij}$$

In Eq. (10.3), $U_1(c_t)q_t^e$ is the utility loss in period t associated with the purchase of an additional unit of the security, while $\delta U_1(c_{t+1})$ measures the units of marginal utility of an additional unit of consumption in period $t+1$ and $(q_{t+1}^e + Y_{t+1})$ is the extra consumption (income) units obtained in period $t+1$ from selling the additional unit of the security in addition to collecting its dividend entitlement. The RHS is thus the expected discounted gain in utility associated with buying the extra unit of the security. The agent is in equilibrium (utility maximizing) at the prevailing price q_t^e if the loss in utility today, which he would incur by buying one more unit of the security $(U_1(c_t)q_t^e)$, is exactly offset by (equals) the expected gain in utility tomorrow $(\delta E_t U_1(\tilde{c}_{t+1})[\tilde{q}_{t+1}^e + \tilde{Y}_{t+1}])$, which the ownership of that additional security will provide. If this equality is not satisfied, the agent will try either to increase or to decrease his holdings of securities.[2]

For the entire economy to be in equilibrium, it must, therefore, be true that:

i. $z_t = z_{t+1} = z_{t+2} = \cdots \equiv 1$, in other words, the representative agent owns the entire security;
ii. $c_t = Y_t$, i.e., ownership of the entire security entitles the agent to all the economy's output;
iii. $U_1(c_t)q_t^e = \delta E_t\{U_1(\tilde{c}_{t+1})(\tilde{q}_{t+1}^e + \tilde{Y}_{t+1})\}$, i.e., the agents' holdings of the security are optimal given the prevailing prices. Substituting (ii) into (iii) informs us that the equilibrium price must satisfy

$$U_1(Y_t)q_t^e = \delta E_t\{U_1(\tilde{Y}_{t+1})(\tilde{q}_{t+1}^e + \tilde{Y}_{t+1})\} \tag{10.4}$$

If there were many firms in this economy—say J firms, with firm j producing the (exogenous) output $\tilde{Y}_{j,t}$—then the same equation would be satisfied for each firm's stock price, $q_{j,t}^e$, i.e.,

$$q_{j,t}^e U_1(c_t) = \delta E_t\{U_1(\tilde{c}_{t+1})(\tilde{q}_{j,t+1}^e + \tilde{Y}_{j,t+1})\} \tag{10.5}$$

where $c_t = \sum_{j=1}^J Y_{j,t}$ in equilibrium.

[2] In equilibrium, however, this is not possible since the supply of securities is fixed. Accordingly, the price will have to adjust until the equality in Eq. (10.3) is satisfied.

Equations (10.4) and (10.5) are the fundamental equations of the CCAPM.[3] A recursive substitution of Eq. (10.2) into itself yields[4]

$$q_t^e = E_t \sum_{\tau=1}^{\infty} \delta^{\tau} \left[\frac{U_1(\tilde{Y}_{t+\tau})}{U_1(Y_t)} \tilde{Y}_{t+\tau} \right] \tag{10.6}$$

establishing the stock price as the sum of all expected discounted future dividends. Equation (10.6) resembles the standard discounting formula of elementary finance, but for the important observation that discounting takes place using the intertemporal marginal rates of substitution defined on the consumption sequence of the representative agent. If the utility function displays risk neutrality and the marginal utility is constant ($U_{11} = 0$), Eq. (10.6) reduces to

$$q_t^e = E_t \sum_{\tau=1}^{\infty} \delta^{\tau} \left[\tilde{Y}_{t+\tau} \right] = E_t \sum_{\tau=1}^{\infty} \left[\frac{\tilde{Y}_{t+\tau}}{(1+r_{\mathrm{f}})^{\tau}} \right] \tag{10.7}$$

which states that the stock price is the sum of expected future dividends discounted at the (constant) risk-free rate. The intuitive link between the discount factor and the risk-free rate leading to the second inequality in Eq. (10.7) will be formally established in Eq. (10.9). The difference between Eqs. (10.6) and (10.7) is the necessity, in a world of risk aversion, of discounting the flow of expected dividends at a rate higher than the risk-free rate, so as to include a risk premium. The question as to the appropriate risk premium constitutes a central issue in financial theory. Equation (10.6) proposes a definite, if not fully operational (due to the difficulty in measuring marginal rates of substitution), answer.

10.3.2 Interpreting the Exchange Equilibrium

To bring about a closer correspondence with traditional asset pricing formulas, we must first relate the asset prices derived previously to rates of return. In particular, we will want to understand, in this model context, what determines the amount by which the risky asset's expected return exceeds that of a risk-free asset. This basic question is also the one for which the standard CAPM provides such a simple, elegant answer ($E\tilde{r}_j = r_{\mathrm{f}} + \beta_j(E\tilde{r}_M - r_{\mathrm{f}})$). Define the period t to $t+1$ return for security j as

$$1 + r_{j,t+1} = \frac{q_{j,t+1} + Y_{j,t+1}}{q_{j,t}}$$

[3] The fact that the representative agent's consumption stream—via his MRS—is critical for asset pricing is true for all versions of this model, including ones with nontrivial production settings. More general versions of this model may not, however, display an identity between consumption and dividends. This will be the case, for example, if the agent receives wage income.

[4] That is, update Eq. (10.4) with q_{t+1}^e on the LHS and q_{t+2}^e in the RHS and substitute the resulting RHS (which now contains a term in q_{t+2}^e) into the original Eq. (10.4); repeat for q_{t+2}^e, q_{t+3}^e, and so on, regroup terms, and extrapolate.

Equation (10.5) may then be rewritten as

$$1 = \delta E_t \left\{ \frac{U_1(\tilde{c}_{t+1})}{U_1(c_t)} (1 + \tilde{r}_{j,t+1}) \right\} \tag{10.8}$$

Let q_t^b denote the price in period t of a one-period riskless discount bond in zero net supply. The bond pays one unit of consumption (income) in every state in the next period. By reasoning analogous to that presented previously,

$$q_t^b U_1(c_t) = \delta E_t\{U_1(\tilde{c}_{t+1})1\}$$

The price q_t^b is the equilibrium price at which the agent desires to hold zero units of the security, and thus supply equals demand. This is so because if he were to buy one unit of this security at a price q_t^b, the loss in utility today would exactly offset the gain in expected utility tomorrow. The representative agent is, therefore, content to hold zero units of the security.

Since the risk-free rate over the period from date t to $t+1$, denoted $r_{\mathrm{f},t+1}$, is defined by $q_t^b(1+r_{\mathrm{f},t+1}) = 1$, we have

$$\frac{1}{1 + r_{f,t+1}} = q_t^b = \delta E_t \left\{ \frac{U_1(\tilde{c}_{t+1})}{U_1(c_t)} \right\}, \tag{10.9}$$

which formally establishes the link between the discount rate and the risk-free rate of return we have used in Eq. (10.7) under the risk neutrality hypothesis. Note that in the latter case ($U_{11} = 0$), Eq. (10.9) implies that the risk-free rate must be a constant.

Now we will combine Eqs. (10.8) and (10.9). Since, for any two random variables $\tilde{x}, \tilde{y}, E(\tilde{x} \cdot \tilde{y}) = E(\tilde{x}) \cdot E(\tilde{y}) + \mathrm{cov}(\tilde{x} \cdot \tilde{y})$, Eq. (10.8) can be written in the form

$$1 = \delta E_t \left\{ \frac{U_1(\tilde{c}_{t+1})}{U_1(c_t)} \right\} E_t\left\{1 + \tilde{r}_{j,t+1}\right\} + \delta \, \mathrm{cov}_t \left\{ \frac{U_1(\tilde{c}_{t+1})}{U_1(c_t)}, \tilde{r}_{j,t+1} \right\} \tag{10.10}$$

Let us make the identification $E_t\{1 + \tilde{r}_{j,t+1}\} = 1 + E(r_{j,t+1})$ while recognizing that the expectation remains conditional on period t information. Substituting Eq. (10.9) into Eq. (10.10) then gives

$$1 = \frac{1 + E\tilde{r}_{j,t+1}}{1 + r_{f,t+1}} + \delta \, \mathrm{cov}_t \left(\frac{U_1(\tilde{c}_{t+1})}{U_1(c_t)}, \tilde{r}_{j,t+1} \right), \text{ or, rearranging,}$$

$$\frac{1 + E\tilde{r}_{j,t+1}}{1 + r_{f,t+1}} = 1 - \delta \, \mathrm{cov}_t \left(\frac{U_1(\tilde{c}_{t+1})}{U_1(c_t)}, \tilde{r}_{j,t+1} \right), \text{ or} \tag{10.11}$$

$$E\tilde{r}_{j,t+1} - r_{f,t+1} = -\delta(1 + r_{f,t+1}) \mathrm{cov}_t \left(\frac{U_1(\tilde{c}_{t+1})}{U_1(c_t)}, \tilde{r}_{j,t+1} \right)$$

Equation (10.11) is the central pricing relationship of the CCAPM, and we must consider its implications. The LHS of Eq. (10.11) is the risk premium on security j. Equation (10.11) tells us that the risk premium on a security will be large when $\text{cov}_t((U_1(\tilde{c}_{t+1}))/(U_1(c_t)), \tilde{r}_{j,t+1})$ is large and negative, i.e., for those securities paying high returns when consumption is high (and thus when $U_1(c_{t+1})$ is low), and low returns when consumption is low (and $U_1(c_{t+1})$ is high). These securities are not very desirable for consumption risk reduction (consumption smoothing): they pay high returns when investors do not need them (consumption is already high) and low returns when they are most needed (consumption is low). Since they are not desirable, they have a low price and thus high expected returns relative to the risk-free security.

The CAPM tells us that a security is relatively undesirable and thus commands a high return when it covaries positively with the market portfolio, i.e., when its return is high precisely in those circumstances when the return on the market portfolio is also high, and conversely. The CCAPM is not in contradiction with this basic idea, but it adds some further degree of precision. From the viewpoint of smoothing *consumption* and risk diversification, an asset is desirable if it has a high return when consumption is low and vice versa.

When the portfolio and asset pricing problem is placed in its proper multiperiod context, the notion of the utility of end-of-period wealth (our paradigm of Chapters 5–8) is no longer relevant, and we have to go back to the more fundamental formulation in terms of the utility derived from consumption, $U(c_t)$. It then becomes clear that the possibility of expressing the objective as maximizing the utility of end-of-period wealth in the two-date setting has, in some sense, lured us down a false trail: fundamentally, the key to an asset's value is its return covariation with the marginal utility of consumption, not with the marginal utility of wealth.

Equation (10.11) has the unappealing feature that the risk premium is defined, in part, in terms of the marginal utility of consumption, which is not observable. To eliminate this feature, we shall make the following approximation.

Let $U(c_t) = ac_t - \frac{b}{2}c_t^2$ (i.e., a quadratic utility function or a truncated Taylor series expansion of a general $U(\ )$), where $a > 0$, $b > 0$, and the usual restrictions apply on the range of consumption. It follows that $U_1(c_t) = a - bc_t$; substituting this into Eq. (10.11) gives

$$
\begin{aligned}
E\tilde{r}_{j,t+1} - r_{\mathrm{f},t+1} &= -\delta(1 + r_{\mathrm{f},t+1})\text{cov}_t\left(\tilde{r}_{j,t+1}, \frac{a - b\tilde{c}_{t+1}}{a - bc_t}\right) \\
&= -\delta(1 + r_{\mathrm{f},t+1})\frac{1}{a - bc_t}\text{cov}_t(\tilde{r}_{j,t+1}, \tilde{c}_{t+1})(-b), \text{ or} \\
E\tilde{r}_{j,t+1} - r_{\mathrm{f},t+1} &= \frac{\delta b(1 + r_{\mathrm{f},t+1})}{a - bc_t}\text{cov}_t(\tilde{r}_{j,t+1}, \tilde{c}_{t+1}).
\end{aligned}
\tag{10.12}
$$

$$E\tilde{r}_{j,t+1} - r_{\mathrm{f},t+1} = \frac{\delta b(1 + r_{\mathrm{f},t+1})}{a - bc_t}\text{cov}_t(\tilde{r}_{j,t+1}, \tilde{c}_{t+1})$$

Equation (10.12) makes our earlier point as to the fundamental source of an asset's value easier to grasp: since the term in front of the covariance expression is necessarily positive, if next-period consumption covaries in a large positive way with $r_{j,t+1}$, then the risk premium on j will be high.

10.3.3 The Formal CCAPM

As a final step in our construction, let us denote the portfolio most highly correlated with consumption by the index $j = c$, and its expected rate of return for the period from t to $t+1$ by $\overline{r}_{c,t+1}$.

Equation (10.12) applies as well to this security, so we have

$$E\tilde{r}_{c,t+1} - r_{\mathrm{f},t+1} = \left[\frac{\delta b(1 + r_{\mathrm{f},t+1})}{a - bc_t}\right]\mathrm{cov}_t(\tilde{r}_{c,t+1}, \tilde{c}_{t+1}) \tag{10.13}$$

Dividing Eq. (10.12) by Eq. (10.13) and thus eliminating the term $[(\delta(1 + r_{\mathrm{f},t+1})b)/(a - bc_t)]$, one obtains

$$\begin{aligned}
\frac{E\tilde{r}_{j,t+1} - r_{\mathrm{f},t+1}}{E\tilde{r}_{c,t+1} - r_{\mathrm{f},t+1}} &= \frac{\mathrm{cov}_t(\tilde{r}_{j,t+1}, \tilde{c}_{t+1})}{\mathrm{cov}_t(\tilde{r}_{c,t+1}, \tilde{c}_{t+1})}, \text{ or} \\
\frac{E\tilde{r}_{j,t+1} - r_{\mathrm{f},t+1}}{E\tilde{r}_{c,t+1} - r_{\mathrm{f},t+1}} &= \frac{\dfrac{\mathrm{cov}_t(\tilde{r}_{j,t+1}, \tilde{c}_{t+1})}{\mathrm{var}(\tilde{c}_{t+1})}}{\dfrac{\mathrm{cov}_t(\tilde{r}_{c,t+1}, \tilde{c}_{t+1})}{\mathrm{var}(\tilde{c}_{t+1})}}, \text{ or} \\
E\tilde{r}_{j,t+1} - r_{\mathrm{f},t+1} &= \frac{\beta_{j,c_t}}{\beta_{c,c_t}}\left[E\tilde{r}_{c,t+1} - r_{\mathrm{f},t+1}\right]
\end{aligned} \tag{10.14}$$

for $\beta_{j,c_t} = (\mathrm{cov}_t(\tilde{r}_{j,t+1}, \tilde{c}_{t+1}))/(\mathrm{var}(\tilde{c}_{t+1}))$, the consumption-$\beta$ of asset j, and $(\mathrm{cov}_t(\tilde{r}_{c,t+1}, \tilde{c}_{t+1}))/(\mathrm{var}(\tilde{c}_{t+1}))$, the consumption-$\beta$ of portfolio c. This equation defines the CCAPM.

If it is possible to construct a portfolio c such that $\beta_{c,c_t} = 1$, the direct analogue to the CAPM is obtained, with $\tilde{r}_{c,t+1}$ replacing the expected return on the market and β_{j,c_t} the relevant beta:

$$E\tilde{r}_{j,t+1} - r_{\mathrm{f},t+1} = \beta_{j,c_t}(E\tilde{r}_{c,t+1} - r_{\mathrm{f},t+1}) \tag{10.15}$$

10.4 Pricing Arrow–Debreu State-Contingent Claims with the CCAPM

Chapter 9 focused on the notion of an Arrow–Debreu state claim as the basic building block for all asset pricing, and it is important to understand what form these securities and their prices assume in the CCAPM setting. Our treatment will be very general and will accommodate more complex settings where the state is characterized by more than one variable.

Whatever model we happen to use, let s_t denote the state in period t. In the prior sections, s_t coincided with the period t output, Y_t.

Given that we are in state s in period t, what is the price of an Arrow–Debreu security that pays one unit of consumption if and only if state s' occurs in period $t+1$? We consider two cases:

1. Let the number of possible states be finite; denote the Arrow–Debreu price as

$$q(s_{t+1} = s'; s_t = s)$$

with the prime superscript referring to the value taken by the random state variable in the next period. Since this security is assumed to be in zero net supply,[5] it must satisfy, in equilibrium,

$$U_1(c(s))q(s_{t+1} = s'; s_t = s) = \delta U_1(c(s'))\,\text{prob}(s_{t+1} = s'; s_t = s), \text{or}$$
$$q(s_{t+1} = s'; s_t = s) = \delta \frac{U_1(c(s'))}{U_1(c(s))}\,\text{prob}(s_{t+1} = s'; s_t = s)$$

As a consequence of our maintained stationarity hypothesis, the same price occurs when the economy is in state s and the claim pays one unit of consumption in the next period if and only if state s' occurs, whatever the current time period t. We may thus drop the time subscript and write

$$q(s'; s) = \delta \frac{U_1(c(s'))}{U_1(c(s))}\,\text{prob}(s'; s) = \delta \frac{U_1(c(s'))}{U_1(c(s))}\pi_{ss'}$$

in the notation of our transition matrix representation. This is Eq. (9.1).

2. For a continuum of possible states, the analogous expression is

$$q(s'; s) = \delta \frac{U_1(c(s'))}{U_1(c(s))} f(s'; s)$$

where $f(s'; s)$ is the conditional density function on s_{t+1} given s, evaluated at s'.

[5] Recall that the very existence of a representative agent required that the underlying multiagent economy possessed a complete financial market structure.

Note that *under risk neutrality*, we have a reconfirmation of our earlier identification of Arrow–Debreu prices as being proportional to the relevant state probabilities, with the proportionality factor corresponding to the time discount coefficient:

$$q(s'; s) = \delta f(s'; s) = \delta \pi_{ss}$$

If these prices are for one-period state-contingent claims, how is an N-period claim priced? They are priced exactly analogously:

$$q^N(s_{t+N} = s'; s_t = s) = \delta^N \frac{U_1(c(s'))}{U_1(c(s))} \text{ prob}(s_{t+N} = s'; s_t = s)$$

The price of an N-period risk-free discount bound $q_t^{b,N}$ given state s is thus given by

$$q_t^{b,N}(s) = \delta^N \sum_{s'} \frac{U_1(c(s'))}{U_1(c(s))} \text{prob}(s_{t+N} = s'; s_t = s) \tag{10.16}$$

or, in the continuum of states notation,

$$q_t^{b,N}(s) = \delta^N \int_{s'} \frac{U_1(c(s'))}{U_1(c(s))} f_N(s'; s) ds' = E_s\left\{ \delta^N \frac{U_1(c_{t+N}(s'))}{U_1(c(s))} \right\}$$

where the expectation is taken over all possible states s' conditional on the current state being s.[6]

Now let us review Eq. (10.6) in the light of the expressions we have just derived.

$$\begin{aligned} q_t^e &= E_t \sum_{\tau=1}^{\infty} \delta^\tau \left[\frac{U_1(\tilde{c}_{t+\tau})}{U_1(c_t)} \tilde{Y}_{t+\tau} \right] \\ &= \sum_{\tau=1}^{\infty} \sum_{s'} \delta^\tau \left[\frac{U_1(c_{t+\tau}(s'))}{U_1(c_t(s))} Y_{t+\tau}(s') \right] \text{prob}(s_{t+\tau} = s'; s_t = s) \\ &= \sum_{\tau} \sum_{s'} q^\tau(s', s) Y_{t+\tau}(s') \end{aligned} \tag{10.17}$$

What this development tells us is that taking the appropriately discounted (at the intertemporal marginal rate of substitution (MRS)) sum of expected future dividends is simply valuing the stream of future dividends at the appropriate Arrow–Debreu prices! The fact that there are no restrictions in the present context in extracting the prices of Arrow–Debreu contingent claims is indicative of the fact that this economy is one of complete markets.

[6] The corresponding state probabilities are given by the Nth power of the matrix $\mathbf{T}$.

Applying the same substitution to Eq. (10.6) as employed to obtain Eq. (10.10) yields

$$
\begin{aligned}
q_t^e &= \sum_{\tau=1}^{\infty} \delta^{\tau} \left\{ E_t\left[\frac{U_1(\tilde{c}_{t+\tau})}{U_1(c_t)}\right] E_t\left[\tilde{Y}_{t+\tau}\right] + \operatorname{cov}\left(\frac{U_1(\tilde{c}_{t+\tau})}{U_1(c_t)}, \tilde{Y}_{t+\tau}\right) \right\} \\
&= \sum_{\tau=1}^{\infty} \delta^{\tau} \left\{ E_t\left[\frac{U_1(\tilde{c}_{t+\tau})}{U_1(c_t)}\right] E_t\left[\tilde{Y}_{t+\tau}\right] \left[1 + \frac{\operatorname{cov}\left(\frac{U_1(\tilde{c}_{t+\tau})}{U_1(c_t)}, \tilde{Y}_{t+\tau}\right)}{E_t\left[\frac{U_1(\tilde{c}_{t+\tau})}{U_1(c_t)}\right] E_t\left[\tilde{Y}_{t+\tau}\right]} \right] \right\}
\end{aligned}
$$

where the expectations operator applies across all possible values of the state output variable, with probabilities given on the line corresponding to the current state s_t in the matrix $\mathbf{T}$ raised to the relevant power (the number of periods to the date of availability of the relevant cash flow).

Using the expression for the price of a risk-free discount bond of τ periods to maturity derived earlier and the fact that $(1+r_{\mathrm{f},t+\tau})^{\tau} q_t^{b,\tau} = 1$ we can rewrite this expression as

$$
q_t^e = \sum_{\tau=1}^{\infty} \frac{\left\{ E_t[Y_{t+\tau}] \left\{ 1 + \frac{\operatorname{cov}(U_1(\tilde{c}_{t+\tau}), \tilde{Y}_{t+\tau})}{E_t[U_1(\tilde{c}_{t+\tau})] E_t[\tilde{Y}_{t+\tau}]} \right\} \right\}}{(1+r_{\mathrm{f},t+\tau})^{\tau}} \tag{10.18}
$$

The quantity being discounted (at the risk-free rate applicable to the relevant period) in the present value term is the equilibrium certainty equivalent of the real cash flow generated by the asset. This is the analogue for the CCAPM of the CAPM expression derived in Section 8.3.

If the cash flows exhibit no stochastic variation (i.e., they are risk free), then Eq. (10.18) reduces to

$$
q_t^e = \sum_{\tau=1}^{\infty} \frac{Y_{t+\tau}}{(1+r_{\mathrm{f},t+\tau})^{\tau}}
$$

This relationship will be derived again in Chapter 12 where we discount risk-free cash flows at the term structure of interest rates. If, on the other hand, the cash flows are risky, yet investors are risk neutral (constant marginal utility of consumption), Eq. (10.18) becomes

$$
q_t^e = \sum_{\tau=1}^{\infty} \frac{E_t[\tilde{Y}_{t+\tau}]}{(1+r_{\mathrm{f},t+\tau})^{\tau}} \tag{10.19}
$$

which is identical to Eq. (10.7) once we recall, from Eq. (10.9), that the risk-free rate must be constant under risk neutrality.

Equation (10.18) is fully in harmony with the intuition of Section 10.3: if the representative agent's consumption is highly positively correlated with the security's real cash flows, the certainty equivalent values of these cash flows will be smaller than their expected values (namely, $\text{cov}(U_1(c_{t+\tau}), Y_{t+\tau}) < 0$). This is so because such a security is not very useful for hedging the agent's future consumption risk. As a result, it will have a low price and a high expected return. In fact, its price will be less than what it would be in an economy of risk-neutral agents (Eq. (10.19)). The opposite is true if the security's cash flows are negatively correlated with the agent's consumption.

10.4.1 The CCAPM and Risk-Neutral Valuation

While the center of our attention in the present chapter is and has been the CCAPM, it is useful to make the connection to Chapter 9's notion of risk-neutral valuation. To see the connection, let $\{\tilde{Y}_t\}$ be an arbitrary uncertain income stream to be priced. Using Eq. (10.17), we can express its price conditional on state s in period t, $q_t\ (s_t = s)$ as

$$\begin{aligned} q_t(s_t = s) &= E_t \sum_{\tau=1}^{\infty} \frac{\delta^\tau U_1(c_{t+\tau}(s'))}{U_1(c_t(s))} Y_{t+\tau}(s_{t+\tau} = s') \\ &= \sum_{\tau=1}^{\infty} \sum_{s'} \frac{\delta^\tau U_1(c_{t+\tau}(s'))}{U_1(c_t(s))} Y_{t+\tau}(s_{t+\tau} = s')\pi(s_{t+\tau} = s';\ s_t = s) \end{aligned} \tag{10.20}$$

From Eq. (10.16), we also know that $(1 + r_{\mathrm{f},t+\tau}(s))^\tau q_t^{b,\tau}(s) = 1$, where $r_{\mathrm{f},t+\tau}(s)$ is the per period rate of return on a risk-free discount bond with price $q_t^{b,\tau}(s)$ paying one unit of the consumption good in every state that may materialize at time $t + \tau$. Accordingly, Eq. (10.20) may be written as

$$\begin{aligned} &q_t(s_t = s) \\ &= \sum_{\tau=1}^{\infty} \sum_{s'} \frac{1}{(1 + r_{\mathrm{f},t+\tau}(s))^\tau} \left[\frac{\dfrac{\delta^\tau U_1(c_{t+\tau}(s'))}{U_1(c_t(s))}}{q_t^{b,\tau}(s)} \right] Y_{t+\tau}(s_{t+\tau} = s')\pi(s_{t+\tau} = s'; s_t = s) \\ &= \sum_{\tau=1}^{\infty} \sum_{s'} \frac{1}{(1 + r_{\mathrm{f},t+\tau}(s))^\tau} \left[\frac{\dfrac{\delta^\tau U_1(c_{t+\tau}(s'))\pi(s_{t+\tau} = s'; s_t = s)}{U_1(c_t(s))}}{\sum_{s'} \dfrac{\delta^\tau U_1(s_{t+\tau} = s')\pi(s_{t+\tau} = s'; s_t = s)}{U_1(c_t(s))}} \right] Y_{t+\tau}(s_{t+\tau} = s') \\ &= \sum_{\tau=1}^{\infty} \sum_{s'} \frac{1}{(1 + r_{\mathrm{f},t+\tau}(s))^\tau} \pi^{RN}(s_{t+\tau} = s'; s_t = s) Y_{t+\tau}(s_{t+\tau} = s') \\ &= \sum_{\tau=1}^{\infty} \frac{1}{(1 + r_{\mathrm{f},t+\tau}(s))^\tau} E_{t+\tau}^{RN} Y_{t+\tau}(s_{t+\tau} = s') \end{aligned} \tag{10.21}$$

As in Chapter 9, the asset is once again priced equal to its expected future cash flows compiled using the risk-neutral probabilities, discounted at the risk-free rate. The applicable risk-free rate is endogenous to the underlying consumption process and may differ for different times to maturity. Note also that

$$\pi^{RN}(s_{t+\tau} = s'; s_t = s) = \left[\frac{\dfrac{U_1(c_{t+\tau}(s'))\pi(s_{t+\tau} = s'; s_t = s)}{U_1(c_t(s))}}{\displaystyle\sum_{s'} \frac{U_1(c_{t+\tau}(s'))\pi(s_{t+\tau} = s'; s_t = s)}{U_1(c_t(s))}} \right]$$

For those future period $t + \tau$ states for which consumption is very low relative to current consumption $c_t\,(s)$, the marginal utility of consumption is high relative to current consumption marginal utility. In this case, the risk-neutral probability of the low consumption ("bad") state is accorded a weight $((U_1(c_{t+\tau}(s')))/(U_1(c_t(s))))$ greater than one relative to the true probability, while correspondingly high consumption states are premultiplied by numbers less than one. It is in this sense that risk-neutral probabilities are "pessimistic" (cf. Section 9.6) relative to the objective state probabilities.

Viewed in this light the CCAPM can be seen as translating the notion of risk-neutral valuation into a setting where there is a direct theoretical connection between the macroeconomy (aggregate consumption) and the financial markets where assets are priced. To evaluate the CCAPM, we will "feed" the observed stochastic process on US consumption into the model and explore whether the model implied asset pricing relationships and derived rates of return reasonably replicate what is observed in actual return data. If they do, we can be more confident that the CCAPM is a reasonable basis for understanding asset pricing relationships.

10.5 Testing the CCAPM: The Equity Premium Puzzle

In this section, we discuss the empirical validity of the CCAPM. Unfortunately, a set of simple and robust empirical observations has been put forward that falsifies this model in an unusually strong way. As a result, we are led to question the model's underlying hypotheses and, a fortiori, those assumptions underlying some of the less sophisticated models seen before. In this instance, the recourse to sophisticated econometrics for drawing significant lessons about our approach to modeling financial markets is superfluous.

A few key empirical observations regarding financial returns in US markets are summarized in Table 10.2, which shows that over a long period of observation the average *ex post* return on a diversified portfolio of US stocks (the market portfolio, as approximated in the United States by the S&P 500) has been close to 7% (in real terms, net of inflation) while the return on 1-year T-bills (taken to represent the return on the risk-free asset) has averaged

Table 10.2: Properties of US asset returns

	US Economy	
	(a)	(b)
R	6.98	16.54
r_f	0.80	5.67
$r - r_f$	6.18	16.67

(a) Annualized mean values in percent.
(b) Annualized standard deviation in percent.
Source*: Data from Mehra and Prescott (1985)*.

less than 1%. These twin observations make up for an equity risk premium of 6.2%. This observation is robust in the sense that it has applied in the United States for a very long period and in several other important countries as well (see Chapter 2). Its meaning is not totally undisputed, however. Goetzmann and Jorion (1999), in particular, argue that the high return premium obtained for holding US equities is the exception rather than the rule.[7]

Here we will take the 6% equity premium at face value, as has much of the huge literature that followed the uncovering of the *equity premium puzzle* by Mehra and Prescott (1985). The puzzle is this: Mehra and Prescott argue that the CCAPM is completely unable, once reasonable parameter values are inserted in the model, to replicate such a high observed equity premium.

Let us illustrate their reasoning.[8] According to the CCAPM, the only factors determining the characteristics of security returns are the representative agent's utility function, his subjective discount factor, and the process on consumption (which equals output or dividends in the exchange economy equilibrium). First, consider the utility function. It is natural in light of the development in Chapter 4 and the requirements for the existence of a representative agent (Box 10.1) to assume that the agent's period utility function displays constant relative risk aversion (CRRA); thus, let us set

$$U(c) = \frac{c^{1-\gamma}}{1-\gamma}$$

[7] Using shorter, mostly postwar, data, premia close to or even higher than the US equity premium are obtained for France, Germany, the Netherlands, Sweden, Switzerland, and the United Kingdom (see, for example, Campbell, 1998). Goetzmann and Jorion, 1999, however, argue that such data samples do not correct for crashes and period of market interruptions, often associated with World War II and thus are not immune to survivorship bias. To correct for such a bias, they assemble long data series for all markets that existed during the twentieth century. They find that the United States has had "by far the highest uninterrupted real rate of appreciation of all countries, at about 5% annually. For other countries, the median appreciation rate is about 1.5%.

[8] The discussion that follows in this section is based on work by Rajnish Mehra.

Empirical studies associated with this model have placed γ in the range of (1, 2). A convenient consequence of this utility specification is that the intertemporal MRS can be written as

$$\frac{U_1(c_{t+1})}{U_1(c_t)} = \left(\frac{c_{t+1}}{c_t}\right)^{-\gamma} \tag{10.22}$$

The second major ingredient is the consumption process. In our version of the model, consumption is a stationary process: It does not grow through time. In reality, however, consumption *is* growing through time. In a growing economy, the analogous notion to the variability of consumption is variability in the growth rate of consumption.

Let $g_{t+1} = (c_{t+1}/c_t)$ denote per capita consumption growth, and assume, for illustration, that g_t is independently and identically lognormally distributed through time. For the period 1889 through 1978, the US economy aggregate consumption has been growing at an average rate of 1.83% annually, with a standard deviation of 3.57% and a slightly negative measure of autocorrelation (−0.14) (cf. Mehra and Prescott, 1985).

The remaining item is the agent's subjective discount factor δ: What value should it assume? Time impatience requires, of course, that $\delta < 1$, but this is insufficiently precise. One logical route to its estimation is as follows: Roughly speaking, the equity in the CCAPM economy represents a claim to the aggregate income from the underlying economy's entire capital stock. We have just seen that, in the United States, equity claims to private capital flows average a 7% annual real return, while debt claims average 1%.[9] Furthermore, the economywide debt-to-equity rates are not very different from 1. These facts together suggest an overall average real annual return to capital of about 4%.

If there were no uncertainty in the model, and if the constant growth rate of consumption were to equal its long-run historical average (1.0183), the asset pricing Eq. (10.8) would reduce to

$$1 = \delta E_t\left\{\left(\frac{\tilde{c}_{t+1}}{c_t}\right)^{-\gamma}\tilde{R}_{t+1}\right\} = \delta(\overline{g})^{-\gamma}\overline{R} \tag{10.23}$$

where $\tilde{R}_{t+1}$ is the gross rate of return on capital and the upper bars denote historical averages.[10] For $\gamma = 1, \overline{g} = 1.0183$, and $\overline{R} = 1.04$, we can solve for the implied δ to obtain $\delta \cong 0.97$. Since we have used an annual estimate for $\overline{g}$, the resulting δ must be viewed as an annual or yearly subjective discount factor; on a quarterly basis it corresponds to $\delta \cong 0.99$. If, on the other hand, we want to assume $\gamma = 2$, Eq. (10.19) solves for $\delta \cong 0.99$ on an annual basis, yielding a quarterly δ even closer to 1. This reasoning demonstrates that assuming

[9] Strictly speaking, these are the returns to publicly traded debt and equity claims. If private capital earns substantially different returns, however, capital is being inefficiently allocated; we assume this is not the case.

[10] Time averages and expected values should coincide in a stationary model, provided the time series is of sufficient length.

higher rates of risk aversion would be incompatible with maintaining the hypothesis of a time discount factor less than 1. While in the case of positive consumption growth, we could technically entertain the possibility of a negative rate of time preference, and thus of a discount factor larger than 1, we rule it out on grounds of plausibility.

At the root of this difficulty is the low return on the risk-free asset (1%), which will haunt us in other ways. As we know, highly risk-averse individuals want to smooth consumption over time, meaning they want to transfer consumption from good times to bad times. When consumption is growing predictably, the good times lie in the future. Agents want to borrow now against their future income. In a representative agent model, it is difficult to reconcile growing consumption with a low rate on borrowing: everyone is on the same side of the market, a fact that inevitably forces a higher rate. This problem calls for an independent explanation for the abnormally low average risk-free rate (e.g., in terms of the liquidity advantage of short-term government debt as in Bansal and Coleman, 1996) or the acceptance of the possibility of a negative rate of time preference so that future consumption is given more weight than present consumption. We will not follow either of these routes here, but rather will, in the course of the present exercise, limit the coefficient of relative risk aversion to a maximum value of 2.

With these added assumptions we can manipulate the fundamental asset pricing Eq. (10.3) to yield two equations that can be used indirectly to test the model. The key step in the reasoning is to demonstrate that, in the context of these assumptions, the equity price formula takes the form

$$q_t^e = vY_t$$

where v is a constant coefficient. That is, the stock price at date t is proportional to the dividend paid at date t (Box 10.2).[11] To confirm this statement, we use a standard trick consisting of guessing that this is the form taken by the equilibrium pricing function, and then verifying that this guess is indeed borne out by the structure of the model. Under the $q_t^e = vY_t$ hypothesis, Eq. (10.3) becomes

$$vY_t = \delta E_t \left\{ (v\tilde{Y}_{t+1} + \tilde{Y}_{t+1}) \frac{U_1(\tilde{c}_{t+1})}{U_1(c_t)} \right\}$$

Using Eq. (10.22) and dropping the conditional expectations operator, since $\tilde{g}_t$ is independently and identically distributed (i.i.d.) through time (its mean is independent of time), we can rewrite this equation as

$$v = \delta E \left\{ (v+1) \frac{\tilde{Y}_{t+1}}{Y_t} \tilde{g}_{t+1}^{-\gamma} \right\}$$

[11] Note that this property holds true as well for the example developed in Box 10.2 as Eqs. (iv) and (v) attest.

BOX 10.2 Calculating the Equilibrium Price Function

Equation (10.4) implicitly defines the equilibrium price series. Can it be solved directly to produce the actual equilibrium prices $\{q(Y^j): j = 1,2, \ldots, N\}$? The answer is positive. First, we must specify parameter values and functional forms. In particular, we need to select values for δ and for the various output levels Y^j, to specify the probability transition matrix $\mathbf{T}$ and the form of the representative agent's period utility function (a CRRA function of the form $U(c) = (c^{1-\gamma})/(1-\gamma)$ is a natural choice). We may then proceed as follows.

Solve for the $\{q(Y^j): j = 1,2, \ldots, N\}$ as the solution to a system of linear equations. Note that Eq. (10.4) can be written as the following system of linear equations (one for each of the N possible current states Y^j):

$$U_1(Y^1)q(Y^1) = \delta\sum_{j=1}^{N}\pi_{1j}U_1(Y^j)Y^j + \delta\sum_{j=1}^{N}\pi_{1j}U_1(Y^j)q(Y^j)$$

$$\vdots \qquad \vdots$$

$$U_1(Y^N)q(Y^N) = \delta\sum_{j=1}^{N}\pi_{Nj}U_1(Y^j)Y^j + \delta\sum_{j=1}^{N}\pi_{Nj}U_1(Y^j)q(Y^j)$$

with unknowns $q(Y^1), q(Y^2), \ldots, q(Y^N)$. Note that for each of these equations, the first term on the RHS is simply a number, while the second term is a linear combination of the $q(Y^j)$s. Barring a very unusual output process, this system will have a solution: one price for each Y^j, i.e., the equilibrium price function.

Let us illustrate: Suppose $U(c) = \ln(c)$, $\delta = 0.96$ and $(Y^1, Y^2, Y^3) = (1.5, 1, 0.5)$—an exaggeration of *boom*, *normal*, and *depression* times. The transition matrix is taken to be as found in Table 10.3.

The equilibrium conditions implicit in Eq. (10.4) then reduce to

$$\begin{aligned} Y^1 &: \tfrac{2}{3}q(1.5) = 0.96 + 0.96\left\{\tfrac{1}{3}q(1.5) + \tfrac{1}{4}q(1) + \tfrac{1}{2}q(0.5)\right\} \\ Y^2 &: q(1) = 0.96 + 0.96\left\{\tfrac{1}{6}q(1.5) + \tfrac{1}{2}q(1) + \tfrac{1}{2}q(0.5)\right\} \\ Y^3 &: 2q(0.5) = 0.96 + 0.96\left\{\tfrac{1}{6}q(1.5) + \tfrac{1}{4}q(1) + 1q(0.5)\right\} \end{aligned} \tag{i}$$

or,

Table 10.3: Transition matrix

$$\begin{array}{c|c} & \begin{matrix} 1.5 & 1 & 0.5 \end{matrix} \\ \begin{matrix} 1.5 \\ 1 \\ 0.5 \end{matrix} & \begin{bmatrix} 0.5 & 0.25 & 0.25 \\ 0.25 & 0.5 & 0.25 \\ 0.25 & 0.25 & 0.5 \end{bmatrix} \end{array}$$

(Continued)

BOX 10.2 Calculating the Equilibrium Price Function (Continued)

$$Y^1: 0 = 0.96 - 0.347q(1.5) + 0.24q(1) + 0.48q(0.5)$$

$$Y^2: 0 = 0.96 + 0.16q(1.5) - 0.52q(1) + 0.48q(0.5) \quad \text{(ii)}$$

$$Y^3: 0 = 0.96 + 0.16q(1.5) + 0.24q(1) - 1.04q(0.5) \quad \text{(iii)}$$

$$\text{(i)} - \text{(ii) yields}: q(1.5) = \tfrac{0.76}{0.507}q(1) = 1.5q(1) \quad \text{(iv)}$$

$$\text{(ii)} - \text{(iii) gives}: q(0.5) = \tfrac{0.76}{1.52}q(1) = 1/2q(1) \quad \text{(v)}$$

substituting Eqs. (iv) and (v) into Eq. (i) to solve for $q(1)$ yields $q(1) = 24$; $q(1.5) = 36$ and $q(0.5) = 12$ follow.

The market-clearing condition implies that $(Y_{t+1}/Y_t) = g_{t+1}$, thus

$$v = \delta E\{(v+1)\tilde{g}^{1-\gamma}\} = \frac{\delta E\{\tilde{g}^{1-\gamma}\}}{1 - \delta E\{\tilde{g}^{1-\gamma}\}}$$

This is indeed a constant, and our initial guess is thus confirmed!

Taking advantage of the validated pricing hypothesis, we can rewrite the equity return as

$$\tilde{R}_{t+1} \equiv 1 + \tilde{r}_{t+1} = \frac{\tilde{q}^e_{t+1} + \tilde{Y}_{t+1}}{q^e_t} = \frac{v+1}{v}\frac{\tilde{Y}_{t+1}}{Y_t} = \frac{v+1}{v}\tilde{g}_{t+1}$$

Taking expectations, we obtain

$$E_t(\tilde{R}_{t+1}) = E(\tilde{R}_{t+1}) = \frac{v+1}{v}E(\tilde{g}_{t+1}) = \frac{E(\tilde{g})}{\delta E\{\tilde{g}^{1-\gamma}\}}$$

The risk-free rate is (Eq. (10.9))

$$R_{\mathrm{f},t+1} \equiv \frac{1}{q^b_t} = \left[\delta E_t\left\{\frac{U_1(\tilde{c}_{t+1})}{U_1(c_t)}\right\}\right]^{-1} = \frac{1}{\delta}\frac{1}{E\{\tilde{g}^{-\gamma}\}} \quad (10.24)$$

which is seen to be constant under our current hypotheses.

Taking advantage of the lognormality assumption, we can express the ratio of the two preceding equations as (see Appendix 10.2 for details)

$$\frac{E(\tilde{R}_{t+1})}{R_\mathrm{f}} = \frac{E\{\tilde{g}\}E\{\tilde{g}^{-\gamma}\}}{E\{\tilde{g}^{1-\gamma}\}} = \exp\left[\gamma\sigma_g^2\right] \quad (10.25)$$

where σ_x^2 is the variance of $\ln x$. Taking logs, we finally obtain

$$\ln(ER) - \ln(R_\mathrm{f}) = \gamma\sigma_g^2 \tag{10.26}$$

Now, we are in a position to confront the model with the data. Let us start with Eq. (10.26). Feeding in the return characteristics of the US economy and solving for γ, we obtain (see Appendix 10.2 for the computation of σ_g^2),

$$\frac{\ln(ER) - \ln(E_{r\mathrm{f}})}{\sigma_x^2} = \frac{1.0698 - 1.008}{(0.0357)^2} = 50.24 = \gamma$$

Alternatively, if we assume $\gamma = 2$ and multiply by σ_g^2 as per Eq. (10.26), one obtains an equity premium of

$$2(0.00123) = 0.002 = (\ln(ER) - \ln(ER_\mathrm{f}) \cong ER - ER_\mathrm{f} \tag{10.27}$$

In either case, this reasoning identifies a major discrepancy between model prediction and reality. The observed equity premium can only be explained by assuming an extremely high coefficient of relative risk aversion ($\cong$50), one that is completely at variance with independent estimates. An agent with risk aversion of this level would be too fearful to take a bath (many accidents involve falling in a bathtub!) or to cross the street. On the other hand, insisting on a more reasonable coefficient of risk aversion of 2 leads to predicting a minuscule premium of 0.2%, much below the 6.2% that has been historically observed over long periods.[12]

Similarly, it is shown in Appendix 10.2 that $E\{g_t^{-\gamma}\} = 0.97$ for $\gamma = 2$; Eq. (10.24) and the observed value for R_f (1.008) then implies that δ should be larger than 1 (1.02). This problem was to be anticipated from our discussion of the calibration of δ, which was based on reasoning similar to that underlying Eq. (10.24). Here the problem is compounded by the fact that we are using an even lower risk-free rate (0.8%) rather than the steady-state rate of return on capital of 4% used in the prior reasoning. In the present context, this difficulty in calibrating δ or, equivalently, in explaining the low rate of return on the risk-free asset has been dubbed the *risk-free rate puzzle* by Weil (1989). As noted previously, we read this

[12] Azeredo (2012) argues that in the pre-1929 period of the Mehra and Prescott (1985) data set, per capita consumption was mismeasured in the standard reported statistics. If it is measured in a way that is more in keeping with practice in the post-1929 period, the serial correlation in consumption growth over the entire sample period (1889–1978) is 0.42, not −0.14 as Mehra and Prescott (1985) report. When their model is recalibrated to incorporate this high, positive serial correlation in consumption, the model-generated equity premium turns progressively more negative as the CRRAs increase in excess of 2.2. High persistent consumption growth causes even a severely risk averse investor to view the equity security as less risky—and more valuable—than the risk-free one. Once again it is not "cash flow" variation *per se* that measures risk but the pattern of that variation relative to an investor's consumption.

result as calling for a specific explanation for the observed low return on the risk-free asset, one that the CCAPM is not designed to provide.

10.6 Testing the CCAPM: Hansen–Jagannathan Bounds

Another, parallel perspective on the puzzle is provided by the Hansen–Jagannathan (1991) bound. The idea is very similar to our prior test, and the end result is the same. The underlying reasoning, however, postpones as long as possible making specific modeling assumptions. It is thus more general than a test of a specific version of the CCAPM. The bound proposed by Hansen and Jagannathan potentially applies to other asset pricing formulations and it similarly leads to a falsification of the standard CCAPM.

The reasoning goes as follows: Let q_t denote the price of an arbitrary asset. For all homogeneous agent economies, the fundamental equilibrium asset pricing Eq. (10.3) can be expressed as

$$q(s_t) = E_t[m_{t+1}(\tilde{s}_{t+1})X_{t+1}(\tilde{s}_{t+1}); s_t] \tag{10.28}$$

where s_t is the state today (it may be today's output in the context of a simple exchange economy, or it may be something more elaborate as in the case of a production economy), $X_{t+1}(\tilde{s}_{t+1})$ is the total (consumption) return in the next period to the asset owner (e.g., in the case of an exchange economy this equals $(\tilde{q}_{t+1} + \tilde{Y}_{t+1})$) and $m_{t+1}(\tilde{s}_{t+1})$ is the *equilibrium pricing kernel*, also known as the *stochastic discount factor* (SDF):

$$m_{t+1}(\tilde{s}_{t+1}) = \frac{\delta U_1(c_{t+1}(\tilde{s}_{t+1}))}{U_1(c_t)}$$

As before $U_1(\ )$ is the marginal utility of the representative agent and c_t is his equilibrium consumption. Equation (10.28) is thus the general statement that the price of an asset today must equal the expectation of its total payout tomorrow multiplied by the appropriate pricing kernel. For notational simplicity, let us suppress the state dependence, leaving it as understood, and write Eq. (10.28) as

$$q_t = E_t[\tilde{m}_{t+1}\tilde{X}_{t+1}] \tag{10.29i}$$

This is equivalent to

$$1 = E_t[\tilde{m}_{t+1}\tilde{R}_{t+1}] \tag{10.29ii}$$

where $\tilde{R}_{t+1}$ is the gross rate of return to ownership of the asset. Since Eq. (10.29ii) holds for each state s_t, it also holds unconditionally; we thus can also write

$$1 = E[\tilde{m}\tilde{R}]$$

where E denotes the unconditional expectation. For any two assets i and j (to be viewed shortly as the return on the market portfolio and the risk-free return, respectively) it must, therefore, be the case that

$$E[\tilde{m}(\tilde{R}_i - \tilde{R}_j)] = 0, \text{ or}$$
$$E[\tilde{m}\tilde{R}_{i-j}] = 0$$

where, again for notational convenience, we substitute $\tilde{R}_{i-j}$ for $\tilde{R}_i - \tilde{R}_j$. This latter expression furthermore implies the following series of relationships:

$$\begin{aligned} E\tilde{m}E\tilde{R}_{i-j} + \operatorname{cov}(\tilde{m}, \tilde{R}_{i-j}) &= 0, \text{ or} \\ E\tilde{m}E\tilde{R}_{i-j} + \rho(\tilde{m}, \tilde{R}_{i-j})\sigma_m \sigma_{R_{i-j}} &= 0, \text{ or} \\ \frac{E\tilde{R}_{i-j}}{\sigma_{R_{i-j}}} + \rho(\tilde{m}, \tilde{R}_{i-j})\frac{\sigma_m}{E\tilde{m}} &= 0, \text{ or} \\ \frac{E\tilde{R}_{i-j}}{\sigma_{R_{i-j}}} &= -\rho(\tilde{m}, \tilde{R}_{i-j})\frac{\sigma_m}{E\tilde{m}} \end{aligned} \tag{10.30}$$

It follows from Eq. (10.30) and the fact that a correlation is never larger than 1 that

$$\frac{\sigma_m}{E\tilde{m}} > \frac{|E\tilde{R}_{i-j}|}{\sigma_{R_{i-j}}} \tag{10.31}$$

The inequality in expression (10.31) is referred to as the Hansen–Jagannathan lower bound on the pricing kernel. If, as noted earlier, we designate asset i as the market portfolio and asset j as the risk-free return, then the data from Table 10.2 and Eq. (10.31) together imply (for the US economy):

$$\frac{\sigma_m}{E\tilde{m}} > \frac{|E(\tilde{r}_M - r_{\mathrm{f}})|}{\sigma_{r_M - \mathrm{rf}}} = \frac{0.062}{0.167} = 0.37$$

Let us check whether this bound is satisfied for our model. From Eq. (10.22), $\tilde{m}(\tilde{c}_{t+1,}c_t) = \delta(g_{t+1})^{-\gamma}$, the expectation of which can be computed (see Appendix 10.2) to be

$$E\tilde{m} = \delta \exp\left(-\gamma\mu_g + \frac{1}{2}\gamma^2\sigma_g^2\right) = 0.99(0.967945) = 0.96 \text{ for } \gamma = 2$$

In fact, Eq. (10.28) reminds us that Em is simply the expected value of the price of a one-period risk-free discount bound, which cannot be very far away from 1. This implies that for the Hansen–Jagannathan bound to be satisfied, the standard deviation of the pricing kernel cannot be much lower than 0.3; given the information we have on $\tilde{g}_t$, it is a short step to estimate this parameter numerically under the assumption of lognormality. When we perform the calculation (again, see Appendix 10.2), we obtain an estimate for $\sigma_m = 0.002$, which is an order of magnitude lower than what is required for Eq. (10.31) to be satisfied. The message is that it is very difficult to get the equilibrium pricing kernel volatility to be anywhere near the required level. In a homogeneous agent, complete market model with standard preferences, where the variation in equilibrium consumption matches the data, consumption is just too smooth. As a result, the marginal utility of consumption does not vary sufficiently to satisfy the bound implied by the data unless the curvature of the utility function—the degree of risk aversion—is assumed to be astronomically high, an assumption which, as we have seen, raises problems of its own.

10.7 The SDF in Greater Generality

The notion of a stochastic discount factor/pricing kernel $\tilde{m}_t$ turns out to be a very general idea with the form $\delta((U_1(c_{t+1}(\tilde{s}_{t+1})))/(U_1(c_t(s_t))))$ being only a very high-profile, special case with a particularly well-defined economic story. The more general perspective may be expressed as follows: For a sufficiently rich set of security payoffs $\tilde{X}_t$ and their associated prices $q(\tilde{X}_t)$, there always exists a unique pricing kernel $\tilde{m}_t$, such that

$q(\tilde{X}_t) = E\tilde{m}_{t+1}\tilde{X}_{t+1}$. In the present section, we propose to introduce the idea and illustrate its use. For simplicity of presentation our focus will be on pricing assets with one-period payoffs and we drop the time subscript.

Payoffs $\tilde{X}$ are random variables defined on some probability space of events $\mathscr{P}$. Let the set of eligible security payoffs χ be defined as:

$$\chi = \{\tilde{X}: E\tilde{X}^2 < \infty, \tilde{X} \text{ a random variable defined on } \mathscr{P}\}$$

with the expectations operator E defined with respect to the probability space $\mathscr{P}$.

The set χ has the property that if $\tilde{X}$ and $\tilde{Z}$ are two distinct asset payoffs in χ, and if a and b are any two real numbers, then $a\tilde{X} + b\tilde{Z} \in \chi$. With this property, χ is said to constitute a linear space.

Now consider an arbitrary pricing function $q: \chi \mapsto R$ such that for any payoff $\tilde{X} \in \chi$, $q(\ )$ assigns a price to this payoff, $q(\tilde{X})$.[13] It is natural to assume that $q(\ )$ respects the law of one price (LOP), which is simply to say that $q(\ )$ assigns only one price to each $\tilde{X} \in \chi$. An absence of arbitrage opportunities is sufficient for this to be true. We are then led to the following straightforward preliminary result.

[13] The expression "pricing function" is just "formalism": $q(\tilde{X})$ is simply the observed price in the securities markets of the asset with claim to $(\tilde{X})$.

Theorem 10.1 The LOP is satisfied for a pricing function $q(\ ):\chi \mapsto R$ if and only if $q(\ )$ is a linear function defined on X.

Proof $\Rightarrow$ Consider two distinct payoffs $\tilde{X}$ and $\tilde{Z}$ in χ with corresponding prices $q(\tilde{X})$ and $q(\tilde{Z})$. Since χ is a linear space $\tilde{W} \equiv a\tilde{X} + b\tilde{Z} \in \chi$ for any real numbers a and b. Let $q(\tilde{W})$ be the price of $\tilde{W}$.

The payoff $\tilde{W}$ can be created by purchasing a units of the payoff $\tilde{X}$ and b units of the payoff $\tilde{Z}$, with a price of $aq(\tilde{X}) + bq(\tilde{Z})$. Since $q(\ )$ respects the LOP, it must be that $q(\tilde{W}) = aq(\tilde{W}) + bq(\tilde{Z})$. We conclude that the pricing function $q(\ )$ is linear on χ.

$\Leftarrow$ Suppose that $q(\ )$ is a linear pricing function on χ, but that there exists an $\tilde{X} \in \chi$ for which $q(\tilde{X}) = q_1$ and $q(\tilde{X}) = q_2$. By the linearity property of $q(\ )$, $q_1 = q(\tilde{X}) = q\left(\frac{1}{2}\tilde{X} + \frac{1}{2}\tilde{X}\right) = \frac{1}{2}q(\tilde{X}) + \frac{1}{2}q(\tilde{X}) = \frac{1}{2}q_1 + \frac{1}{2}q_2$. It follows from the string of equalities that $q_1 = q_2$ and the LOP holds under the linear pricing function $q(\ )$.

The small observation made in Theorem 10.1 has very large implications when seen as a part of Theorem 10.2.

Theorem 10.2 For any linear pricing function $q(\ ):\chi \mapsto R$, there is a unique $m^* \in \chi$ such that $q(\tilde{X}) = E\tilde{m}^*\tilde{X}$ for all $\tilde{X} \in \chi$. Furthermore, if there are no arbitrage opportunities under $q(\ )$, $m^* > 0$. If the underlying financial market is complete m^* is unique.

Proof Application of the Reiz representation theorem; the proof and examples are found in the Web Notes to this chapter.

Theorem 10.2 merits attention for a number of subtle reasons. First, it is very general: there is no mention of a representative agent whose "preference identity" is not easy to come by (the data certainly suggests that it is not captured by the utility function $U(c) = (c^{1-\gamma})/(1-\gamma)$). All that we know is that the pricing kernel m^* is of the same form as the asset payoffs: if X is the set of normally distributed payoffs, m^* will assume the form of a normal distribution. More generally, there is no specific equilibrium model underlying the conclusion to Theorem 10.2, a feature that is both a strength and a weakness.

It is a strength because it suggests that the overall general equilibrium pricing perspective adopted in this chapter is a reasonable one since it leads to a pricing relationship that must exist theoretically. The obvious downside to Theorem 10.2 is that it is totally uninformative as to how the economic structures (preferences, technologies, and markets) give rise to the asset prices we see arising out of the financial markets. In this sense, it does not contribute to the goal of this text!

Second, as we will see in later chapters, there are useful contexts where the relevant m^* can be precisely calculated from no arbitrage relationships, without reference to an economic model. Finance professionals welcome this feature. Note that the identification of the pricing kernel m^* of Theorem 10.2 is essentially the same as identifying the appropriate risk-neutral probabilities or Arrow–Debreu state prices.

There is one refinement of the conclusion to Theorem 10.2 which concerns the form of the postulated m^* and recalls the by-now-customary notion of an efficient portfolio. We present this refinement in Corollary 10.1.

Corollary 10.1 Consider the pricing kernel m^* of Theorem 10.2. Then there exist associated (with m^*) constants a and b and an associated MV-efficient portfolio $p(m^*)$ such that

$$m^* = a + br_{p(m^*)}$$

where $r_{p(m^*)}$ is the return on that portfolio.

The proof of Corollary 10.1 is constructive and not immediately intuitive so we stop the discussion at this point and invite the reader to refer to the Web Notes where the proof is taken up in detail.[14]

10.8 Some Extensions

10.8.1 Reviewing the Diagnosis

Our first dynamic general equilibrium model thus fails when confronted with actual data. Let us review the source of this failure. Recall our original pricing Eq. (10.9), specialized for a single asset, the market portfolio:

$$\begin{aligned}\bar{r}_{M,t+1} - r_{f,t+1} &= \delta(1 + r_{f,t+1})\mathrm{cov}_t\left(\frac{U_1(\tilde{c}_{t+1})}{U_1(c_t)}, \tilde{r}_{M,t+1}\right) \\ &= -\,\delta(1 + r_{f,t+1})\rho\left(\frac{U_1(\tilde{c}_{t+1})}{U_1(c_t)}, \tilde{r}_{m,t+1}\right)\sigma\left(\frac{U_1(\tilde{c}_{t+1})}{U_1(c_t)}\right)\sigma(\tilde{r}_{M,t+1}) \\ &= -\,(1 + r_{f,t+1})\rho(\tilde{m}_t, \tilde{r}_{M,t+1})\sigma(\tilde{m}_t)\sigma(\tilde{r}_{M,t+1})\end{aligned}$$

Written in this way, it is clear that the equity premium depends upon the standard deviation of the MRS (or, equivalently, the stochastic discount factor), the standard deviation of the return on the market portfolio, and the correlation between these quantities. For the United States, and most other industrial countries, the problem with a model in which pricing and return relationships depend so much on consumption (and thus MRS) variation is that average per

[14] The statement of the Corollary becomes a bit more believable if we write $q(\tilde{X}) = E\tilde{m}^*\tilde{X}, \tilde{X} \in \chi$ as $1 = E\tilde{m}^*(\tilde{X}/q(\tilde{X})) = E\tilde{m}^*\tilde{R}_{\tilde{X}}$ where $\tilde{R}_{\tilde{X}}$ is the gross return on the asset with payoff $\tilde{X} \in \chi$. Written this way, with $\tilde{R}_{\tilde{X}}$ a return function, so also must $\tilde{m}^*$ be a return function by Theorem 10.2.

capita consumption does not vary much at all. If this model is to have any hope of matching the data, we must modify it in a way that will increase the standard deviation of the relevant MRS, or the variability of the dividend being priced (and thus the $\sigma(r_{M,t+1})$). We do not have complete freedom over this latter quantity, however, as it must be matched to the data as well.

These thoughts suggest a number of "strategies" for resolving the joint equity premium and risk-free rate puzzles:

1. Modify the representative agent's utility function in such a way that he is much more sensitive to consumption variation than is evident in the CRRA specification. Most contemporary theories go down this route in some way. We are reminded, however, the Constantinides' (1982) construction of a representative agent presumes that investor preferences are VNM-expected utility, a requirement that constrains "creativity" along the preference dimension.
2. Uncover some feature of the consumption growth process which, although consistent with the basic consumption growth data in Section 10.5, is nevertheless highly objectionable to the representative agent, thus requiring an offsetting additional risk premium on any asset whose payout is closely aligned with this feature. It is "additional" in the sense that it goes beyond the premium for quarterly aggregate dividend risk.
3. Argue that it is not the "representative agent" who bears stock market risk, but the fraction of the population that owns stocks and actually trades them. Government data reveals that equity ownership is highly concentrated: for the US 20% of the population owned 92% of all stock in the year 2010. It is thus the consumption and income processes of this group alone that should be relevant to CCAPM modeling.

To date, most efforts to resolve the equity premium cum risk-free rate puzzles involve some combination of the first and second strategies. At the same time, the set of time series statistics that a model is challenged to explain has expanded. Presently, any CCAPM-style model, if it is to be accorded much credibility, must explain not only the mean market equity return and risk-free rate (and thus also the equity premium), but also the volatilities of these return series as well. Furthermore, all these statistics must be replicated in an environment where the mean and SD of the growth rates of consumption, $\tilde{g}$, and dividends, $\tilde{g}_{\mathrm{dir}}$, also well approximate their empirical counterparts. Using data from Bansal and Yaron (2004), whose work we will review shortly, Table 10.4 refreshes our recollection as to the magnitudes of the relevant quantities involved.

The list now includes such other concerns as demonstrating predictability patterns in stock return data.[15]

[15] The basic predictability phenomenon is most simply laid out in Cochrane (2011). It is the observation that regressions of the form $R_{t+j} = \alpha_j + \beta_j(D_t/P_t) + \varepsilon_j$ have substantial R^2 and highly significant β_j coefficients for $j = 5$. The expression (D_t/P_t) is the market aggregate dividend/price ratio in period t and R_{t+j} is the cumulative market return over the subsequent j periods. It is in the context of regressions of this type that dividend price ratios are said to predict returns.

Table 10.4: Financial statistics[a]

	Mean	SD
r^e	7.19	19.42
r^f	0.86	0.97
r^P	6.33	19.2
G	1.8	2.93
g_{div}	1.8	11.49
$(P/D)_t$	15.5[b]	25.56

[a]Source: Bansal and Yaron (2004); based on data for the period 1928–1998. All numbers measured in percent, annualized.
[b]For the period 1871 to the present time (March 5, 2014) the average (P/D) ratio for the S&P$_{500}$ (or its predecessors) is the figure indicated. As of this writing (March 5, 2014) the actual (P/D) ratio is 19.46.

In the remaining sections of this chapter, we review a number of modeling approaches to explaining at least the basic return statistics of Table 10.4[16] Most are straightforward adaptations of the basic Lucas (1978) tree model of Section 10.3 but generalized to reflect the observed growth rates of consumption and dividends (see Appendix 10.1). Explaining the sources of these most basic return regularities is important not only for progress in financial economics, but also in business cycle and growth theory which are based on the same CCAPM paradigm enriched with a production sector so that the equilibrium consumption/dividend series become endogenous to the model.

10.8.2 Adding a Disaster State

Here the story goes back to Reitz (1988), but remains, in its various manifestations, under active consideration. The notion is as follows: suppose there is a very low probability state of the world (e.g., the Great Depression of the 1930s—a once per 100 years event) in which consumption growth (simultaneously dividend and output growth) turns negative. Such an event will have enormous consequences for investor welfare, ceteris paribus, since *it will lead to lower future consumption levels in all future periods.* Accordingly, equity securities, whose returns are highly correlated with consumption growth, will be very unattractive to investors and must pay high average returns if investors are to be persuaded to hold them. So the logic goes. Adding a low probability disaster state is one example of strategy (2).

To explore the power of this idea Reitz (1988) proposed a very modest generalization of the original Mehra and Prescott (1985) model to accommodate a "disaster state." In particular,

[16] See Lengwiler (2004) for a complementary, and very thorough, discussion of the earlier literature. See also Mehra (2012) for a review of more recent developments.

he describes the evolution of consumption growth by a *three*-state Markov transition matrix of the form:

$$\begin{array}{ll} & \begin{array}{ccc} \quad g_1 \quad & \quad g_2 \quad & g_3 \end{array} \\ \begin{array}{l} g_1 = 1 + \overline{g} + \sigma_g \\ g_2 = 1 + \overline{g} - \sigma_g \\ g_3 = \dfrac{1 + \overline{g}}{2} \end{array} & \begin{bmatrix} \phi & 1 - \phi - \eta & \eta \\ 1 - \phi - \eta & \phi & \eta \\ 1/2 & 1/2 & 0 \end{bmatrix} = T_1 \end{array}$$

where η is the probability of entering the disaster growth state g_3 from either of the "normal" growth states (g_1, g_2). There is equal probability (1/2) of exiting the disaster state to either of the normal states, and zero probability of persisting in it. With $\eta = 0.003$, $\phi = 0.47$, $\gamma = 5.3$, $\delta = 0.98$, $\overline{g} = 0.018$ and $\sigma_g = 0.036$, e.g., Reitz (1988) obtains $Er^e = 6.15\%$ and $Er_f = 0.89\%$ (he does not report σ_{r^e} or σ_{r_f}). With this set of parameters, the long-run likelihood of being in a disaster state is 0.0029, which is small. With consumption falling to roughly one-half its prior level when the disaster ensues, however, 38 years of average growth are required for consumption to recover only to its precrash level. This is a severe disaster indeed.[17,18]

This result is suggestive of the power of the "disaster scenario." It is also an attractive generalization of Mehra and Prescott (1985), since it is entirely consistent with the complete markets-expected utility-representative agent modeling perspective underlying the entirety of this chapter. The subsequent literature is thus principally focused on its

[17] As Reitz (1988) notes, per capita consumption in the United States "only" fell to 78% of its prior level from the period 1929–1933. With $g_3 = 0.75$, he needs a higher disaster probability ($\eta = .008$), higher CRRA ($\gamma = 10$, $\delta = 0.992$, all other parameters unchanged from the case above) to achieve the following results: $Er^e = 6.37\%$, and $Er_f = 2.97\%$. While the average return on equity is about unchanged from the prior example, the risk-free security is in much less relative demand. As a result, its equilibrium price is lower. This outcome is to be expected: the disaster-induced consumption decline is so much less and, despite the increased frequency of entering the disaster state, its long-run likelihood remains less than 1%.

[18] The notion of a "peso problem" is closely related to the concept of a disaster state. A peso problem arises when security returns reflect the possibility of a disaster event even though one has not yet occurred. Suppose a long time series of dividend (equivalently, consumption) growth rates was generated using the Reitz (1988) transition matrix specification T_1, yet it just so happened that the series displayed no occasions of the disaster state (a low-probability event, certainly, but not an impossible one). A representative agent knowing the form T_1 would price the dividend stream (in the manner of Eq. (10.6)) accordingly to reflect the disaster possibility. To an outside observer with the same utility function as the agent but with only information from the observed time series to guide his calculations, it would appear that the equilibrium price was too low. This (seeming) mispricing phenomenon is referred to as a "peso problem." The name comes from a set of events in the period July 1974–July 1976 where the futures price of the Mexican peso in terms of US dollars indicated the possibility of a peso devaluation even though none had been observed. The Mexican peso was subsequently devalued relative to the US dollar in August 1976. For a simple discussion of "peso problem" implications for asset pricing, see Danthine and Donaldson (1999).

calibration: Are there actual events representable by a Reitz (1988) disaster; how does the severity of actual disasters compare to the Reitz (1988) parameterizations, what is the comparative duration, and what is the nature of the relative recoveries (note that a Reitz, 1988 disaster is short-lived)? Mehra and Prescott (1988), in particular, find the relevant Reitz (1988) calibrations basically implausible.

Barro (2006) uses the "disaster experiences" of a sample of 35 countries in the 20th century to calibrate a generalized version of the Reitz (1988) model. Within this sample of countries, he finds 60 occasions where GDP per capita fell by at least 15% which becomes his criterion for a "disaster" (the maximum decline occurred in Germany, 64%, during the 1944–1946 period), from which he estimates a constant disaster probability of 1.7% per year. In the Reitz (1988) formulation, this would amount to a $\eta = 0.017$. When a disaster occurs, Barro (2006) assumes, however, that the extent of the disaster (the percentage output decline) is governed by a probability distribution calibrated to match the historical frequency of disaster severities, with a mean value of 0.29. Barro (2006) also allows for a partial bond default, which is modeled as an event that only occurs conditional on a disaster experience with conditional probability of 0.4. The conditional severity of the bondholders' haircut is assumed to have the same probability density as the disaster severity itself. Basically, Barro (2006) has less severe "consumption disasters" than Reitz (1988), but they are more frequent and affect both bond and stock returns. Under his calibration, Barro (2006) is able to match the equity premium, though at the cost of somewhat excessive risk-free rate estimates. He does not report the volatilities of returns or consumption growth, however.

Barro's (2006) paper rehabilitated the "disaster scenario" of Reitz (1988) by presenting a model calibration that is more empirically plausible, and most of the subsequent "disaster" literature has been focused on generalizing his model in various ways. Gabaix (2012) adds the feature that both the severity of the disaster and the expected recovery rates are themselves time varying, thereby introducing another source of risk. Note that in Reitz (1988) both the disaster severity (g_3) and recovery rate (immediate) are fixed; in Barrow (2006), the disaster severity is a fixed probability distribution and recovery is also immediate. Gabaix (2012) is able to match a very wide class of both equity regularities (e.g., the premium, the SD of the price dividend ratio) and regularities in the bond market.

Nakamura et al. (2013) also takes off from Barro (2006) along similar dimensions: (1) they assume the representative agent's preferences are Epstein–Zin (1989)—see Section 5; and they allow (2) disasters and (3) recoveries to evolve over multiple periods (recall that in Reitz, 1988, and Barro, 2006, disasters and recoveries are both of one-period duration). Essentially their focus is on a richer and more realistic evolution of per capita consumption over disasters and recoveries. In their baseline estimation, they are able to achieve a 4.8% premium and a risk-free rate of 0.10% with an estimated CRRA of 6.4. Volatilities are not reported, however.

10.8.3 Habit Formation

Another device for resolving the equity premium puzzle has been to admit utility functions that exhibit higher rates of risk aversion at the margin and thus can translate small variations in consumption into a large variability of the pricing kernel.[19] One way to achieve this objective without being confronted with the risk-free rate puzzle—which is exacerbated if we simply decide to postulate a higher γ—is to admit some form of *habit formation*. This is the notion that the agent's utility today is determined not only by her absolute consumption level, but also by the relative position of her current consumption vis-à-vis what can be viewed as a *stock of habit*, the latter summarizing either her past consumption history (with more or less weight placed on more distant consumption levels) or the history of aggregate consumption (summarizing in a sense the consumption habits of her neighbors, a "keeping up with the Joneses" effect: see Abel, 1990). This modeling perspective takes the view that an investor's utility of consumption is principally affected by departures from his prior consumption history, either his own or that of a social reference group, i.e., departures from what he may have been accustomed to consume or what he may hold as a socially accepted level of consumption. This concept is open to a variety of different specifications, with diverse implications for behavior and asset pricing. The interested reader is invited to consult Campbell and Cochrane (1999) for a review. Here we will be content to illustrate the underlying working principle. To that end, we specify the representative agent's period preference ordering to be of the form

$$U(c_t, c_{t-1}) \equiv \frac{(c_t - \chi c_{t-1})^{1-\gamma}}{1-\gamma}$$

where $\chi \le 1$ is a parameter. In an extreme case, $\chi = 1$, the period utility depends only upon the deviation of current period t consumption from the prior period's consumption. As we noted earlier, actual data indicate that per capita consumption for the United States and most other developed countries is very smooth. This implies that (c_t/c_{t-1}) is likely to be close to 1 much of the time. For this specification, the agent's effective (marginal) relative risk aversion reduces to $R_R(c_t) = (\gamma/(1-(c_{t-1}/c_t)))$; with $c_t \approx c_{t-1}$, the effective $R_R(c)$ will thus be very high, even with a low γ, and the representative agent will behave as though he is very risk averse to consumption variation. With a careful choice of the habit specification, the risk-free rate puzzle can be avoided (see Constantinides, 1990; Campbell and Cochrane, 1999), and the equity premium increased.

[19] There is a large literature which seeks to find preference representations more amenable to the resolution of the "puzzles" than the basic CRRA family, and we cannot detail that enormous literature here. The interested reader is referred to Donaldson and Mehra (2008) for a reasonably comprehensive review. Many of the proposed orderings are vulnerable to the same criticisms as will be leveled at "habit-formation preferences."

"Habit-formation preferences" are by now a standard feature of many finance and macrofinance models. Its drawbacks as a modeling device are also well known. In some models, the marginal CRRA exceeds 100 in order for a good match to the basic stylized facts to be achieved, a figure that seems excessive especially for wealthy investors who own a preponderance of outstanding stock.[20] In addition, if the growth rate of consumption is modeled as an i.i.d. process, a reasonable first approximation to the data, habit formation can lead to an excessively volatile equilibrium risk-free rate. Campbell and Cochrane (1999) are able to deal with this later drawback but in a habit model variant in which it, curiously, can pay to destroy consumption: the loss in utility during the period in which the consumption is destroyed is more than offset by the enhanced future utility due to a reduced habit! (see Ljungqvist and Uhlig, 2009). In more general models where consumption is endogenous, habit-formation preferences often end up generating equilibrium consumption sequences which are too smooth relative to data. Perhaps the most basic criticism of the habit construct is that to date we lack any choice-theoretic axiomatic foundations for habit formation within the domain of VNM-expected utility preferences, a fact that may mean they are inconsistent with Constantides's (1982) representative agent construct. In this sense, habit-formation preferences are essentially "behavioral" in nature, and the need for an underlying theory is recognized. See Rozen (2010) and Chetty and Szeidl (2004) for progress in this direction. Habit-formation preferences are an illustration of strategy (1) for resolving the puzzle.

10.8.4 The CCAPM with Epstein–Zin Utility

At this stage it is interesting to inquire whether, in addition to its intellectual appeal on grounds of generality, Epstein and Zin's (1989) separation of time and risk preferences might contribute a solution to the equity premium puzzle and more generally, alter our vision of the CCAPM and its message. The work by Nakamura et al. (2013) mentioned earlier suggests this possibility.

Let us start by looking specifically at the equity premium puzzle. It will facilitate our discussion to repeat Eqs. (5.14) and (5.15) defining the Epstein–Zin preference representation (refer to Chapter 5 for a discussion and for the log case):

$$U(c_t, CE_{t+1}) = [(1-\delta){c_t}^{1-\rho} + \delta CE_{t+1}^{1-\rho}]^{\frac{1}{1-\rho}}, \text{where}$$
$$[CE(\tilde{U}_{t+1})]^{1-\gamma} = E_t(\tilde{U}_{t+1})^{1-\gamma}$$

[20] In the United States during the 1990s households in the top 20% of the wealth distribution own 98% of all outstanding stocks.

Weil (1989) uses these preferences in a setting otherwise identical to that of Mehra and Prescott (1985). Asset prices and returns are computed similarly. What he finds, however, is that this greater generality, *per se*, does not resolve the *equity premium puzzle*, but rather tends to underscore what we have already introduced as the *risk-free rate puzzle*.

The Epstein–Zin (1989, 1991) preference representation does not really innovate along the risk dimension, with the parameter γ alone capturing risk aversion in a manner very similar to the standard case. It is, therefore, not surprising that Weil (1989) finds that only if this parameter is fixed at implausibly high levels ($\gamma \approx 45$) can a properly calibrated model replicate the premium—the Mehra and Prescott (1985) result in a different setting. With respect to time preferences, if ρ is calibrated to respect empirical studies, ($\frac{1}{\rho}$, the intertemporal elasticity of substitution is estimated to be about one), then the model also predicts a risk-free rate that is much too high. The reason for this is the same as the one outlined at the end of Section 10.5: separately calibrating the intertemporal substitution parameter ρ tends to strengthen the assumption that the representative agent is highly desirous of a smooth intertemporal consumption stream. With consumption growing on average at 1.8% per year, the agent must be offered a very high risk-free rate in order to be induced to save more, thus making his consumption tomorrow even more in excess of what it is today (less smoothing).

Although Epstein and Zin preferences do not help solve the equity premium puzzle, it is interesting to study a version of the CCAPM with these generalized preferences. The setting is once again a Lucas (1978) style economy with N assets, with the return on the equilibrium portfolio of all assets representing the return on the market portfolio. Using an elaborate dynamic programming argument, Epstein and Zin (1989, 1991) derive an asset pricing equation of the form

$$E_t\left\{\left[\delta\left(\frac{\tilde{c}_{t+1}}{c_t}\right)^{-\rho}\right]^{\theta}\left[\frac{1}{1+\tilde{r}_{M,t+1}}\right]^{1-\theta}(1+\tilde{r}_{j,t+1})\right\} \equiv 1 \tag{10.32}$$

where $\tilde{r}_{M,t}$ denotes the period t return on the market portfolio, $r_{j,t}$ the period t return on some asset in it, and $\theta = ((1-\gamma)/(1-\rho)), 0 < \delta < 1, 1 \neq \gamma > 0, \rho > 0$. Note that when time and risk preferences coincide ($\gamma = \rho, \theta = 1$), Eq. (10.32) reduces to the pricing equation of the standard time-separable CCAPM case.

The pricing kernel itself is of the form

$$\left[\delta\left(\frac{\tilde{c}_{t+1}}{c_t}\right)^{-\rho}\right]^{\theta}\left[\frac{1}{1+\tilde{r}_{M,t+1}}\right]^{1-\theta} \tag{10.33}$$

which is a geometric average (with weights θ and $1-\theta$, respectively) of the pricing kernel of the standard CCAPM, $[\delta(\tilde{c}_{t+1}/c_t)^{-\rho}]$, and the pricing kernel for the $\log(\rho = 0)$ case, $[1/(1+\tilde{r}_{M,t+1})]$.

Epstein and Zin (1991) next consider a linear approximation to the geometric average in Eq. (10.33),

$$\theta\left[\delta\left(\frac{\tilde{c}_{t+1}}{c_t}\right)^{-\rho}\right]+(1-\theta)\left[\frac{1}{1+\tilde{r}_{M,t+1}}\right] \tag{10.34}$$

Substituting Eq. (10.34) into Eq. (10.32) gives

$$\begin{gathered}E_t\left\{\left\{\theta\left[\delta\left(\frac{\tilde{c}_{t+1}}{c_t}\right)^{-\rho}\right]+(1-\theta)\left[\frac{1}{1+\tilde{r}_{M,t+1}}\right]\right\}(1+\tilde{r}_{j,t+1})\right\}\approx 1,\ \text{or}\\ E_t\left\{\theta\left[\delta\left(\frac{\tilde{c}_{t+1}}{c_t}\right)^{-\rho}\right](1+\tilde{r}_{j,t+1})+(1-\theta)\left[\frac{1}{1+\tilde{r}_{M,t+1}}\right](1+\tilde{r}_{j,t+1})\right\}\approx 1\end{gathered} \tag{10.35}$$

Equation (10.35) is revealing. As we noted earlier, the standard CAPM relates the (essential, nondiversifiable) risk of an asset to the covariance of its returns with $\tilde{r}_M$, while the CCAPM relates its riskiness to the covariance of its returns with the growth rate of consumption. With separate time and risk preferences, Eq. (10.35) suggests that both covariances matter for an asset's return pattern.[21] But why do these covariance effects enter separately? The second term leads to the covariance of an asset's return with M which captures its atemporal, nondiversifiable risk (as in the static CAPM model). The first term leads to a covariance of the asset's return with the growth rate of consumption which captures its risk across successive time periods. When risk and time preferences are separated, it is not surprising that both sources of risk should be individually present and that there should exist risk premia associated with each. This relationship is more striking if we assume joint lognormality and heteroskedasticity in consumption and asset returns. Campbell et al. (1997) are able to express Eq. (10.35) in a form whereby the risk premium on asset j satisfies:

$$E_t(\tilde{r}_{j,t+1})-r_{\mathrm{f},t+1}=\delta\frac{\sigma_{j,c}}{\psi}+(1-\delta)\sigma_{j,M}-\frac{\sigma_j^2}{2} \tag{10.36}$$

where $\sigma_{i,c}=\mathrm{cov}(\tilde{r}_{j,t},\tilde{c}_t/(c_{t-1}))$, and $\sigma_{j,M}=cov(\tilde{r}_{j,t},\tilde{r}_{M,t})$. Both sources of risk are clearly present.

There are two important applications of Epstein–Zin (1989) to which we now turn.

[21] To see this, recall that for two random variables $\tilde{x}$ and $\tilde{y}$, $E(\tilde{x}\tilde{y})=E(\tilde{x})E(\tilde{y})+\mathrm{cov}(\tilde{x},\tilde{y})$, and employ this substitution in both terms on the LHS of Eq. (10.35).

10.8.4.1 Bansal and Yaron (2004)

To prepare for a description of the Bansal and Yaron (2004) model, let us return for a moment to the i.i.d. consumption growth paradigm of Section 10.5. It is permissible to express this process as

$$\tilde{g}_t = \overline{g} + \sigma_g \tilde{\varepsilon}_t \tag{10.37}$$

where $\overline{g} = 0.018$, $\sigma_g = 0.036$, and $\{\tilde{\varepsilon}_t\}$ is a sequence of i.i.d. normal random variables with $\{\tilde{\varepsilon}_t\} \sim N(0, 1)$ for all t. Under the Lucas (1978)—Mehra and Prescott (1985) paradigm the growth process on dividends, $\tilde{g}_{\mathrm{div},t}$, is the same: $\tilde{g}_t = \tilde{g}_{\mathrm{div},t}$. With representation (10.37) in mind as our baseline formulation, Bansal and Yaron (2004) propose a generalization with four dimensions.

(i) The stochastic processes on the growth rates of consumption and dividends are specified independently of one another. Among other advantages, this separation allows a greater volatility to be assigned to the dividend series than to the per capita consumption series, as is evident in the data.[22] See Table 10.4.

(ii) Both the dividend growth and consumption growth series share a small, highly persistent long-run component. It is this feature that has led to the sobriquet "long-run risks" model. In other words, the Bansal and Yaron (2004) economy can persist in a regime of low growth realizations for many period (both consumption and dividends), before moving to a high growth regime and vice versa. As such, the growth rate uncertainty (captured, say, by the expected time to transition from one regime to another) is only resolved very slowly, after the passage, on average, of many periods. Since this slowly resolving risk component is shared by both the dividend and consumption series, the covariance of the investor's consumption growth with equity returns can be large.

(iii) The utility specification for the representative agent is Epstein—Zin with parameters $\gamma > 1/\rho$ (risk aversion parameter exceeds the EIS - the elasticity of intertemporal substitution). We recall from our earlier discussion (see Section 5.7.3) that investors with these preference parameters prefer the early resolution of uncertainty. This desired property is exactly what the dividend series of the equity security (the "market portfolio") does *not* provide in the Bansal and Yaron (2004) model because of its long-run risk component.

(iv) Short-run dividend and consumption volatility (these will be risks analogous to the $\sigma_g \tilde{\varepsilon}_t$ term in the baseline representation (10.37)) are time varying. In this way, yet another source of uncertainty is introduced. It is referred to as "stochastic volatility."

[22] In the Bansal and Yaron (2004) model, the difference in these series is labor income.

Taken together, there are three sources of uncertainty faced by an investor participating in the Bansal and Yaron (2004) economy: short-run volatility, long-run volatility, and changing short-run volatility. In equilibrium, each will have an associated risk premium.

To gain a bit more intuition as to the workings of the Bansal and Yaron (2004) model, let us make explicit the growth rate processes discussed above:

$$\tilde{g}_{t+1} = \overline{g} + x_t + \sigma_t \tilde{\eta}_{t+1} \tag{10.38i}$$

$$\tilde{g}_{\text{div},t+1} = \overline{g}_d + \varphi x_t + \varphi_d \sigma_t \tilde{u}_{t+1} \tag{10.38ii}$$

where $\tilde{x}_t$ denotes the long-run risks component and is itself governed by

$$\tilde{x}_{t+1} = \overline{\rho} x_t + \varphi_e \sigma_t \tilde{e}_{t+1} \tag{10.38iii}$$

Note that the choice of autocorrelation coefficient $\overline{\rho}$ determines the persistence of the growth component while the choices of φ and φ_d calibrate (1) the volatility of dividends relative to consumption and (2) their respective correlations with per capita consumption.[23] The shocks $\tilde{\eta}_{t+1}, \tilde{u}_{t+1}$, and $\tilde{\varepsilon}_{t+1}$ are each distributed $N(0, 1)$, and all are assumed to be statistically independent of one another.

Lastly, the volatility of the short-run risks component, σ_t^2, is itself given by a mean reverting stochastic process

$$\tilde{\sigma}_{t+1}^2 = \sigma^2 + v(\sigma_t^2 - \sigma^2) + \sigma_w \tilde{w}_{t+1} \tag{10.38iv}$$

Again, $\tilde{w}_{t+1}$ is i.i.d. $N(0,1)$, while $v > 0$ governs the extent of mean reversion to the long-run average, σ^2.

Although we will not detail the solution technique here, or the approximations that make it feasible, the net effect of the three risk sources is to lead to an equilibrium return expression parallel to (10.36)

$$\begin{aligned} E_t(\tilde{r}_{M,t+1}) - r_{\text{f},t+1} = {} & \beta_{M,\eta} \gamma \sigma_t^2 + \beta_{M,e} \left(\gamma - \frac{1}{\rho} \right) \overline{\text{K}}_1 \sigma_t^2 \\ & + \beta_{M,w} \left(\gamma - \frac{1}{\rho} \right) (1-\gamma) \overline{\text{K}}_2 \sigma_w^2 - \frac{1}{2} \text{var}_t(\tilde{r}_{M,t+1}) \end{aligned} \tag{10.39}$$

where $\overline{K}_1$ and $\overline{K}_2$ are constants determined by $\overline{\rho}$, φ_e, etc. Note that each of the three distinct risk manifestations has its own "designated premium." The model attains an excellent fit to the data, not only insofar as replicating the equity premium and risk-free rate, but also as regards their respective return volatilities and the volatility of the dividend price ratio. The importance

[23] Bansal and Yaron set $\overline{\rho} = 0.979$.

of the long-run risks component $\tilde{x}_t$ is forcefully present by the fact that if $\rho = 0$ (so that consumption and dividend risk become i.i.d.), the premium declines essentially to zero.[24]

The Bansal and Yaron (2004) model is a complex one. Nevertheless, it remains a direct descendant of the basic CCAPM of this chapter. While there are more sources of risk, each one commands a recognizable premium. Agent preferences are hypothesized to create high aversion to the new sources of risk. As such, Bansal and Yaron (2004) follow strategies (1) and (2) presented at the close of Section 10.8.1.

Unfortunately, the long-run risks model of consumption growth is disputed along a number of notable dimensions vis-à-vis the data. Beeler and Campbell (2012), in particular, argue that the long-run risks model, as calibrated in Bansal and Yaron (2004), leads to excessive persistence in consumption and dividend growth and excessive predictability (by equity prices) of these quantities relative to what is found in the data. A subsequent calibration (Bansal et al., 2011) undertaken to resolve the aforementioned shortcomings, while successful in its immediate goal, has the unfortunate implication that the model's equilibrium real term structure is downward sloping, which is also not generally observed. It appears that Bansal and Yaron (2004) may not be the last word on the equity premium puzzle.

We are thus left with two models of the economy, each of which can be argued reflects the economy reasonably well: the traditional i.i.d. consumption growth model and the Bansal and Yaron (2004) model of consumption growth with a long-run risks component. Which of these "world views" should the representative agent−investor adopt? On the one hand, the long-run risks model does a much better job of explaining financial data. On the other hand, the representative agent's welfare—the present value of his discount future expected utility—is much higher under the i.i.d. consumption growth perspective. Under a "robust control" perspective, the representative agent−investor would choose between the two consumption growth models based on the "worst-case" scenario, and thus act as though he believed the long-run risks model was fact.

10.8.4.2 Collin-Dufresne et al. (2013)

Thus far in this chapter, all the models discussed have the feature that the representative agent either knows the true parameter values of the stochastic environment in which he operates or has subjective beliefs as to those values but does not change his beliefs as events unfold. In the standard Lucas (1978)−Mehra and Prescott (1985) paradigm, for example, the representative agent is assumed to know the actual parameters of the consumption growth process, $\overline{g}$ and σ_g. Even in the case where the modeling exercise includes the estimation of model parameters (such as in Nakamura et al., 2013), these very

[24] This result confirms the earlier work of Weil (1989) who argued that Epstein−Zin preferences, *per se*, will not guarantee a resolution of the equity premium and risk-free rate puzzles.

parameters are then held to be fixed and accurately measured by the representative agent as the economy evolves and the model time series of asset prices and returns are generated.

Collin-Dufresne et al. (2013) abandon this perspective and explore the asset pricing implications of allowing the representative agent to learn the pertinent stochastic process parameters on, for example, consumption growth. To illustrate the consequences of this innovative perspective, consider a standard Lucas (1978)–Mehra and Prescott (1985) setting (consumption equals dividends) with consumption growth evolving in the customary way ($\tilde{g}_{t+1} = \overline{g} + \sigma_g \tilde{\varepsilon}_{t+1}$, where $\{\tilde{\varepsilon}_t\}$ is i.i.d., $N(0,1)$) modified by assuming the representative agent does not know $\overline{g}$. Rather, he "begins the day" with a "prior" estimate $\overline{g} \sim N(\overline{g}_0, A_0 \sigma_g^2)$.[25,26]

Collin-Dufresne et al. (2013) assume the agent employs Bayes' rule to update his prior distribution recursively as time passes and he gains additional information from the actual, observed g_t realizations. Accordingly, after t realizations have been observed, his estimate for the distribution governing $\overline{g}$ evolves as per

$$\overline{g} \sim N(\overline{g}_t, A_t \sigma_g^2) \text{ where} \tag{10.39i}$$

$$A_{t+1} = \frac{A_t}{1 + A_t} < A_t, A_0 = 1, \text{ and} \tag{10.39ii}$$

$$\overline{g}_{t+1} = \left(\frac{A_t}{1 + A_t}\right) g_t + \left(\frac{1}{1 + A_t}\right) \overline{g}_t \tag{10.39iii}$$

Note that the updated mean of the subjective distribution is computed as a weighted average of the prior period's estimate and the current period's growth realization.[27]

From the representative agent's (subjective) perspective, the consumption dynamics he faces become

$$\tilde{g}_{t+1} = \overline{g}_t + (\sqrt{1 + A_t}) \sigma_g \tilde{\varepsilon}_{t+1}, \ \tilde{\varepsilon}_{t+1} \sim N(0, 1) \text{ and} \tag{10.40}$$

$$\overline{g}_{t+1} = \overline{g}_t + \frac{A_t}{\sqrt{1 + A_t}} \sigma_g \tilde{\varepsilon}_{t+1} \tag{10.41}$$

The import of relationships (10.40) and (10.41) is that changes (updates) in $\overline{g}_t$ will have permanent effects as regards the agent's continuation utility and thus exercise substantial

---

[25] Our discussion going forward follows directly from Collin-Dufresne et al. (2013).

[26] For simplicity the representative agent is assumed to know σ. We know, at least, that σ can be estimated more precisely relative to the mean.

[27] See the Web Notes on Bayesian updating.

influence on his asset demands and the economy's equilibrium asset prices and returns.[28] Note that changes in the estimated $\overline{g}_t$ become a source of "long-run risk" for the agent. The origins of this risk component, however, are very different from the long-run risk component in Bansal and Yaron (2004). In that latter study, the source of long-run risk lay in the presumed, highly persistent consumption growth component. In Collin-Dufresne et al. (2013), the source of long-run risk lies in the changing estimate of the true mean consumption growth rate.

From our discussion of Bansal and Yaron (2004), we have learned that long-run risks are relevant for asset pricing only when the representative agent dislikes them. Accordingly, the second pillar of the Collin-Dufresne et al. (2013) model is the assumption of Epstein–Zin preferences for the representative agent with $\gamma > \frac{1}{\rho}$: a preference for the early resolution of uncertainty. In the event that $\rho = 1$ (see Eq. (5.14b)), these authors are able to solve directly for the market risk premium:

$$E\tilde{r}_{M,t} - r_{f,t} = \gamma\sigma_g^2 + \left(\frac{\gamma - 1}{\tilde{\delta}}\right)A_t\sigma_g^2 \tag{10.42}$$

where $\tilde{\delta} = -\ln\delta$, $\tilde{\delta}$ is the representative agent's time subjective time preference parameter and γ his Epstein–Zin risk parameter. Note that in Eq. (10.42), the time index t denotes the tth period in the learning process.

If the agent is indifferent to the timing of uncertainty resolution so that $\gamma = \frac{1}{\rho} = 1$, then the long-run risk component of the premium, $((\gamma - 1)/\tilde{\delta})A_t\sigma_g^2$, disappears, and the premium reverts to its value in the standard CRRA utility case (see Eq. (10.26)).

As $t \mapsto \infty$, $A_t \mapsto 0$ and the long-run risks premium gradually disappears. How quickly does it happen? Using parameter values taken from Bansal and Yaron (2004) ($\gamma = 10$, $\delta = 0.994$) together with $A_0 = 1$, Collin-Dufresne et al. (2013) find that the equity risk premium with learning, while it declines rapidly with the passage of time, is still significantly larger than the premium under the known-parameter scenario:

$$\frac{\left(\begin{array}{l}\text{equity risk}\\ \text{premium with}\\ \text{unknown}\\ \text{mean growth}\\ \text{parameter}\end{array}\right)}{\left(\begin{array}{l}\text{equity risk}\\ \text{premium with}\\ \text{known mean}\\ \text{growth}\\ \text{parameter}\end{array}\right)} = \frac{\gamma\sigma_g^2 + \left(\dfrac{\gamma - 1}{-\ln\delta}\right)A_t\sigma_g^2}{\gamma\sigma_g^2} = \begin{cases} 151 & t = 0 \\ 1.37 & t = 100 \text{ years} \\ 1.19 & t = 200 \text{ years} \end{cases}$$

[28] Note that for this system and $A_0 = 1$ (since the agent knows the true SD), $A_t \mapsto 0$ and $\overline{g}_t \mapsto \overline{g}$.

These results suggest that even after 100 years of learning, when the variance of the agent's subjective distribution of the mean growth rate has been substantially diminished, the effect of the remaining uncertainty on the representative agent's continuation utility is still large, with an elevated equity premium as the result.[29]

An especially interesting analysis in Collin-Dufresne et al. (2013) concerns the representative agent learning the probability of the economy's transition to a "disaster state." Here we have a situation where the "disaster effect" and the "parameter-learning effect" can, potentially, work together to resolve the various puzzles.

The basic innovation here is to assume that the true model of consumption growth conforms to a process

$$\tilde{g}_t = \overline{g}(\tilde{s}_t) + \sigma(\tilde{s}_t)\tilde{\varepsilon}_t$$

where the "state," $\tilde{s}_t$, follows a two-state Markov chain parameterized as follows:

$$\begin{array}{cc} & (\overline{g}(s_1),\ \sigma(s_1)) \quad\quad (\overline{g}(s_2),\ \sigma(s_2)) \\ \begin{array}{c}(\overline{g}(s_1),\ \sigma(s_1)) \\ (\overline{g}(s_2), \sigma(s_2))\end{array} & \begin{bmatrix} \pi_{11} & 1-\pi_{11} \\ \pi_{21} & \pi_{22} \end{bmatrix} \end{array}$$

$$\begin{aligned} (\overline{g}(s_1), \sigma(s_1)) &= (0.54\%, 0.98\%) \\ (\overline{g}(s_2), \sigma(s_2)) &= (-1.15\%, 1.47\%) \end{aligned} \qquad \text{(quarterly)}$$

$$\begin{aligned} \pi_{11} &= 0.9975 \\ \pi_{22} &= 0.9325 \end{aligned} \qquad \text{(quarterly)}$$

The disaster event $(\overline{g}(s_2), \overline{\sigma}(s_2))$ is intended to replicate the fact that in the Great Depression, consumption growth declined at an average annual rate of -4.6% from 1929 to 1933 while the transition probabilities are chosen so as to result, on average, in one 4-year depression (a succession of 16 disaster states) per 100-year time horizon. Collin-Dufresne et al. (2013) also break the dividend–output–consumption equivalence of Mehra and Prescott (1985) by postulating an independent process on dividends of the form

[29] What about uncertainty in the σ_g^2 parameter? Collin-Dufresne et al. (2013), page 12) give a succinct answer: "... as it is easier to learn a constant volatility parameter than a mean parameter and since volatility is a second-order effect in terms of utility the asset pricing effects of learning about the variance of shocks (even in the case of Epstein–Zin preferences) are very quickly negligible."

Table 10.5: Summary financial statistics learning model[a]

	Data		Model	
	Mean	SD	Mean	SD[b]
$\tilde{r}^e$	5.96		6.58	
$\tilde{r}_f$	0.86	0.97	0.91	0.63
$\tilde{r}_p$	5.10	20.21	5.67	16.23

[a]Source: Table entries are from Collin-Dufresne et al. (2013), Table 1, Panel A.
[b]SD (r^e) not reported; data value from Bansal and Yaron (2004).

$$g_{\text{div},t} = \overline{g} + \lambda(g - \overline{g}) + \sigma_{\text{div}}\tilde{\eta}_{t+1}$$

where $\{\eta_t\}$ is i.i.d., $N(0, 1)$, and $\text{cov}(\tilde{\eta}_t, \tilde{\varepsilon}_t) = 0$. The parameter λ is included to reflect the fact that dividend growth reacts more dramatically to business cycle variations than consumption growth, while σ_{div} is chosen to match real dividend growth volatility over their data period (11.5%).[30]

The model is then simulated and return statistics based on successively generating 20,000 sample paths of artificial return data, each 400 model periods (100 years for a quarterly calibration) in length. The return statistics reported in Table 10.5 represent averages of the return statistics computed across all of the 20,000 artificially generated sample paths. As part of the data generating strategy, agents learn about the parameters for 400 model periods prior to the initiation of data collection. To be clear, the data from which each sample path's statistics are computed is the equilibrium consequence of agents' continued learning after the completion of a prior 100-year learning period.[31] The results of this extensive exercise represent an excellent replication of the basic financial stylized facts (See Table 10.5).

Many other experiments are reported in this article, with the results largely supportive of their modeling hypothesis.

Let us review where we stand. As with Reitz (1988), the background context is a much-feared disaster state in which the agent's marginal utility of consumption is high: he desperately wishes to consume more. In Reitz (1988), the probability of entering the disaster state is precisely known to the agent, and he exits the state immediately. In Collin-Dufresne et al. (2013), the agent slowly learns the probabilities of entering the disaster state and exiting from it (after, potentially, an extended period of time). Each time period, new

[30] The parameter λ is estimated by regressing annual real dividend growth on annual consumption growth using data from the period 1929–2011. In particular, the authors find $\lambda = 2.5$ and subsequently use this figure. See the next section of this chapter for a partial explanation for the behavior of dividends relative to aggregate consumption.

[31] It is as if the agents learned about the disaster likelihood using nineteenth century economic data before entering the twentieth century.

information is learned (a transition event is observed or not) and the relevant probabilities updated. A changing estimate of the probability of entering the disaster state can have large consequences for the agent's continuation utility because of the permanent, consumption level consequences of altered future growth rate perceptions. As a result, the agent perceives the claim to the dividend stream to be a highly risky asset.

As the agent continues to learn, he gradually becomes aware of the model's exact parameter values, in which case the premium declines to roughly 1%, and the risk-free rate climbs to roughly 3%: in a world with known risks with mean consumption growing, agents need a substantial return to postpone consumption thereby making it less smooth. All in all, this story strikes us as a plausible one.

Looking across all the disaster-related models we have discussed so far, Reitz (1988), Barro (2006), Nakamura et al. (2013), and Collin-Dufresne et al. (2013), the true parameters of the evolution of uncertainty in the model economy are assumed to be known: either they are subjectively held or they are assumed to have been properly estimated from historical data. In the case of Collin-Dufresne et al. (2013), the agent learns the parameter values that are assumed to be known if only to us, the readers: we know exactly what the agent is learning about but he himself does not. What if the parameters (such as σ_g) are not known, but are governed themselves by a hypothetical stochastic process. This issue is dealt with in Weitzman (2007) to follow.

10.8.5 Beyond a Representative Agent and Rational Expectations

10.8.5.1 Beyond a Representative Agent

Another approach to addressing the outstanding financial puzzles focuses on recognizing that only a small fraction of the population holds substantial financial assets, stocks in particular.[15] This fact implies that only the variability of the consumption stream of the stockholding class should matter for pricing risky assets. There are reasons to believe that the consumption patterns of this class of the population are both more variable and more highly correlated with stock returns than average per capita consumption.[32] Observing, furthermore, that wages are very stable and that the aggregate wage share is countercyclical (i.e., proportionately larger in bad times when aggregate income is relatively low), it is reasonable to suggest that firms, and thus their owners, the shareholders, implicitly insure workers against income fluctuations associated with the business cycle. If this is a significant feature of the real world, it should have implications for asset pricing, as we presently demonstrate.

[32] Mankiw and Zeldes (1991) confirm this conjecture. They indeed find that shareholder consumption is 2.5 times as variable as nonshareholder consumption. Data problems, however, preclude taking their results as more than indicative.

Before trying to incorporate such a feature into a CCAPM-type model, it is useful first to recall the notion of risk sharing. Consider the problem of allocating an uncertain income (consumption) stream between two agents so as to maximize overall utility. Assume, furthermore, that these income shares are not fixed across all states but can be allocated on a state-by-state basis. This task can be summarized by the allocation problem

$$\max_{c_1(\theta), c_2(\theta)} U(c_1(\tilde{\theta})) + \lambda V(c_2(\tilde{\theta})), \text{ s.t.}$$
$$c_1(\tilde{\theta}) + c_2(\tilde{\theta}) \leq Y(\tilde{\theta})$$

where $U(\ )$, $V(\ )$ are, respectively, the two agents' utility functions, $c_1(\tilde{\theta})$ and $c_2(\tilde{\theta})$ their respective income assignments, $Y(\tilde{\theta})$ the economy-wide state-dependent aggregate income stream, and λ their relative weight.

The necessary and sufficient first-order condition for this problem is

$$U_1(c_1(\tilde{\theta})) = \mu V_1(c_2(\tilde{\theta})) \tag{10.43}$$

Equation (10.43) states that the ratio of the marginal utilities of the two agents should be constant. We have seen it before as Eq. (8.3). As we saw there, it can be interpreted as an optimal risk-sharing condition in the sense that it implicitly assigns more of the income risk to the less risk-averse (more risk-tolerant) agent. To see this, take the extreme case where one of the agents, say the one with utility function $V(\ )$, is risk neutral—indifferent to risk. According to Eq. (10.43), it will then be optimal for the other agent's income stream to be constant across all states: he will be perfectly insured. Agent $V(\ )$ will thus absorb all the risk (in exchange for a higher average income share).

To understand the potential place of these ideas in the CCAPM setting, let $V(\ )$ now denote the period utility function of the representative shareholder, and $U(\ )$ the period utility function of the representative worker who is assumed to hold no financial assets and who consequently consumes his wage w_t. As before, let Y_t be the uncertain (exogenously given) output. The investment problem of the shareholders—the maximization problem with which we began this chapter—now becomes

$$\max_{\{z_t\}} E\left(\sum_{t=0}^{\infty} \delta^t V(\tilde{c}_t)\right)$$
$$\text{s.t.}$$
$$c_t + p_t z_{t+1} \leq z_t d_t + p_t z_t$$
$$d_t = Y_t - w_t$$
$$U_1(w_t) = \lambda V_1(d_t),$$
$$z_t \leq 1, \ \forall t$$

Here we simply introduce a distinction between the output of the tree, Y_t, and the dividends paid to its owners, d_t, on the plausible grounds that workers need to be paid to take care of the trees and collect the fruits. This payment is w_t. Moreover, we introduce the idea that the wage bill may incorporate a risk insurance component, which we formalize by assuming that the variability of wage payments is determined by an optimal risk sharing rule equivalent to Eq. (10.43). One key parameter is the income share, μ, which may be interpreted as reflecting the relative bargaining strengths of the two groups. Indeed, a larger μ gives more income to the worker.

Assets in this economy are priced as before, with Eq. (10.1) becoming

$$V_1(c_t)p_t = \delta E_t\{V_1(\tilde{c}_{t+1})(\tilde{p}_{t+1} + \tilde{d}_{t+1})\} \tag{10.44}$$

While the differences between Eqs. (10.1) and (10.38) may appear purely notational, their importance cannot be overstated. First, the pricing kernel derived from Eq. (10.44) will build on the firm owners' MRS, defined over shareholder consumption (dividend) growth rather than the growth in average per capita consumption. Moreover, the definition of dividends as output minus a stabilized stream of wage payments opens up the possibility that the flow of payments to which firm owners are entitled is effectively not only much more variable than consumption but more variable than output as well. Therein lies a concept of leverage, one that has been dubbed *operating leverage*, similar to the familiar notion of financial leverage. In the same way that bondholders come first and are entitled to a fixed, noncontingent interest payment, workers also have priority claims to the income stream of the firm, and macroeconomic data on the cyclical behavior of the wage share confirm that wage payments are more stable than aggregate income.

These ideas are most simply laid out in Danthine and Donaldson (2002) to which the reader is referred for details, and we find that this class of models can generate significantly increased equity premia. When an extra notion of distributional risk, associated with the possibility that μ varies stochastically, is added, in a way that permits better accounting of the observed behavior of the wage share over the medium run, Danthine and Donaldson (2002) find the premium approaches 6%, a fact that is not entirely surprising since a fundamentally new source of risk with its own premium is being introduced. Favilukis and Lin (2013) argue, in a production economy context, that it is impossible to replicate the financial stylized facts without an operating leverage component, while Santos and Veronesi (2006) demonstrate that the ratio of labor income to consumption is helpful in explaining long-horizon stock returns.

Aside from the worker-firm owner dichotomy, heterogeneous agent models have generally not been prominent in the equity premium resolution literature.[33]

[33] An exception to this statement is Dumas (1989) who considers an economy with two agents of differing CRRAs. In equilibrium, the less risk averse agent owns a larger fraction of the wealth than in bad times. The corresponding (wealth distribution dependent) representative agent exhibits time varying risk aversion. This feature also improves model performance.

10.8.5.2 Beyond Rational Expectations

Let us review the modeling philosophy that has governed this chapter's discussion thus far. In the case of Mehra and Prescott (1985), Barro (2006), Bansal and Yaron (2004), and most others, the representative agent is assumed to know the identity of the stochastic process governing the evolution of the uncertain quantities of relevance to him and the precise parameters governing that process. While the parameters may be the result of an estimation procedure (e.g., Nakamura et al., 2013), the representative agent in the model is presumed to take the values on faith. The case of Collin-Dufresne et al. (2013) is a bit different but in the same spirit: the model builder is presumed to know the relevant model parameters about which the representative agent learns as the model's time path evolves. This overall perspective is a reflection of the idea that the economy is stationary and has been operating for a very long time, sufficiently long for economic participants to have learned all the relevant stochastic parameters. As a modeling philosophy, it is loosely labeled as the "rational expectations" view of the world.

Weitzman (2007) forcefully disputes this view of the world, arguing that there will always exist "fundamental structural uncertainty" associated, in particular, with the consumption growth rate parameter σ_g. Suppose that $\tilde{\sigma}_{g,t}$ is time dependent and evolves stochastically in a nonergodic way (it does not converge to a fixed σ_g^*), and that the representative agent employs Bayesian methods to learn about $\tilde{\sigma}_{g,t}$. Weitzman (2007) points out that the Bayesian learning will lead to fat tailed predictive student t probability distributions governing consumption growth. Under this probability distribution, Geweke (2001) has shown earlier that expected utility may not exist for standard CRRA utility, a tendency that must be held in check by postulating a reasonable prior distribution at the start of the learning process. Weitzman (2007) refers to his perspective as the "Bayesian evolutionary-learning thickened-tail explanation" of the equity premium. Weitzman (2007) argues that standard methodologies for learning in an evolutionary context lead to a reversal of the standard baseline financial anomalies. The model derived equity premium will be too large, the equity return volatility too large, and the risk-free rate too low (potentially negative) relative to data. In Weitzman (2007), the puzzle is not that the observed equity premium is as large as it is, but that it is so small relative to what it might be under his Bayesian evolutionary learning perspective. We close with a series of quotes taken from Weitzman's paper:

> *Weitzman (2007)... "formalizes the idea that non ergodic parameter uncertainty leads to permanently tail thickened distribution of growth rates that can cause expected marginal utility to blow up—and shows a rigorous series in which 'containing the student-*t *explosion' necessitates an unavoidable dependence of asset prices upon some form or another of exogenously imposed subjective beliefs."*

> *"This potentially explosive outcome remains the mathematical driving force behind the scene, which imparts the statistical illusion of an enormous equity premium incompatible*

with the standard neoclassical paradigm. When people are peering forward into the future they are also looking backward at their own prior and what they are seeing there is a spooky reflection of their own present insecurity in not being able to judge accurately the possibility of bad evolutionary mutations of future history t might conceivably ruin equity investors by wiping out their stock market holdings at a time t when their world has already taken a very bad turn."

"... for asset pricing applications it is not at all unscientific to adhere to the non REE idea that no amount of past data can be... large enough to identify the relevant structural uncertainty concerning future economic growth. Moreover, as a corollary, REE calibrations ignoring this basic principle of learning about hidden evolutionary parameters may very well end up badly underestimating the comparative utility risk of a real-world gamble in the unknown structural potential economic growth, relative to a nearly safe investment in a nearly sure thing."

10.9 Conclusions

The various modifications considered in the previous sections represent many of the most recent contributions to the equity premium literature. As is apparent, there is no consensus and active research continues. Nevertheless, the overall equilibrium stochastic discount factor seems secure because of its natural connection to the data. Even Weitzman (2007) does not dispute the overall research program.

At this juncture, one may nevertheless be led to the view that structural asset pricing theory, based on rigorous dynamic general equilibrium models, provides limited operational support in our quest for the understanding of time series financial market phenomena. This state of affairs perhaps explains the popularity of less encompassing approaches based on the concept of arbitrage to be reviewed in succeeding chapters.

References

Abel, A., 1990. Asset prices under habit formation and catching up with the joneses. Am. Econ. Rev. 80, 38−42.

Azeredo, F., 2012. The equity premium: a deeper puzzle. Discussion Paper, Navigant Economics.

Bansal, R., Coleman, W.J., 1996. A monetary explanation of the equity premium, term premium and risk-free rates puzzles. J. Polit. Econ. 104, 1135−1171.

Bansal, R., Yaron, A., 2004. Risks for the long run: a potential resolution of asset pricing puzzles. J. Finan. 59, 1481−1509.

Barro, R., 2006. Rare disasters and asset markets in the twentieth century. Q. J. Econ. 121, 823−866.

Barro, R.J., 1974. Are government bonds net wealth? J. Polit. Econ. 82, 1095−1117.

Bansal, J., Kiku, D., Yaron, A., 2011. An empirical evaluation of the long-run risks model of asset prices. Working Paper, Duke University and the University of Pennsylvania.

Beeler, J., Campbell, J., 2012. The long-run risks model and aggregate asset prices: an empirical assessment. Crit. Finan. Rev. 1, 141−182.

Campbell, J., Lo, A., MacKinlay, A.C., 1997. The econometrics of financial markets. Princeton University Press, Princeton, NJ.

Campbell, J.Y., 1998. Asset prices, consumption, and the business cycle. NBER Working Paper 6485, March 1998, forthcoming in the *Handbook of Macroeconomics*, Amsterdam, North Holland.

Campbell, J.Y., Cochrane, J.H., 1999. By force of habit: a consumption-based explanation of aggregate stock market behavior. J. Polit. Econ. 107, 205–251.

Chetty, R., Szeidl, A., 2004. Consumption commitments: neoclassical foundations for habit formation. NBER Working Paper #10970.

Collin-Dufresne, P., Johannes, M., Lochstoer, L., 2013. Parameter learning in general equilibrium. Working Paper, Columbia University, Graduate School of Business.

Constantinides, G.M., 1982. Intertemporal asset pricing with heterogeneous consumers and without demand aggregation. J. Bus. 55, 253–267.

Constantinides, G.M., 1990. Habit formation: a resolution of the equity premium puzzle. J. Polit. Econ. 98, 519–543.

Danthine, J.P., Donaldson, J., 2002. Labor relations and asset returns. Rev. Econ. Stud. 69, 41–64.

Danthine, J.P., Donaldson, J., 1999. Non-falsified expectations and general equilibrium asset pricing. Econ. J. 109, 607–635.

Donaldson, J., Mehra, R., 2008. Risk-based explanations of the equity premium. Handbook of the Equity Premium. Elsevier, New York and Amsterdam.

Dumas, B., 1989. Two person dynamic equilibrium in the capital market. Rev. Finan. Stud. 2, 157–188.

Epstein, L., Zin, S., 1989. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: a theoretical framework. Econometrica. 57, 937–969.

Epstein, L., Zin, S., 1991. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: an empirical analysis. J. Polit. Econ. 99, 263–286.

Favilukis, J., Lin, X., 2013. Wage rigidity: a quantitative solution to several asset pricing puzzles. Working Paper, London School of Economics and Ohio State University.

Gabaix, X., 2012. Variable rare disasters: an exactly solved framework for ten puzzles in macro-finance. Q. J. Econ. 127, 645–700.

Geweke, J., 2001. A note on some limitations of CRRA utility. Econ. Lett. 71, 341–345.

Goetzmann, W., Jorion, P., 1999. A century of global stock markets. J. Finan. 55, 953–980.

Guvenen, F., 2011. Macroeconomics with heterogeneity: a practical guide. Working Paper, University of Minnesota.

Hansen, L., Jagannathan, R., 1991. Implications of security market data for models of dynamic economies. J. Polit. Econ. 99, 225–262.

Lengwiler, Y., 2004. Microfoundations of Financial Economics. Princeton University Press, Princeton, NJ.

Lucas, R.E., 1978. Asset pricing in an exchange economy. Econometrica. 46, 1429–1445.

Ljungqvist, L., Uhlig, H., 2009. Optimal endowment destruction under Campbell–Cochrane. NBER Working Paper #14772.

Mankiw, G., Zeldes, S., 1991. The consumption of stockholders and non-stockholders. J. Finan. Econ. 29, 97–112.

Mehra, R., 2012. Consumption-based asset pricing models. Annu. Rev. Financ. Econ. 4, 385–409.

Mehra, R., Prescott, E.C., 1985. The equity premium: a puzzle. J. Monet. Econ. 15, 145–161.

Mehra, R., Prescott, E.C., 1988. The equity risk premium: a solution? J. Monet. Econ. 22, 133–136.

Nakamura, E., Steinsson, J., Barro, R., Ursúa, J., 2013. Crises and recoveries in an empirical model of consumption disasters. Am. Econ. J.: Macroecon. 5, 35–74.

Reitz, T., 1988. The equity premium: a solution. J. Monet. Econ. 22, 117–131.

Rozen, K., 2010. Foundations of intrinsic habit formation. Econometrica. 78, 1341–1373.

Rubinstein, M., 1974. An aggregation theorem for securities markets. J. Finan. Econ. 1, 225–244.

Santos, T., Veronesi, P., 2006. Labor income and predictable stock returns. Rev. Econ. Stud. 19, 1–44.

Weil, P.h., 1989. The equity premium puzzle and the risk-free rate puzzle. J. Monet. Econ. 24, 401–421.

Weitzman, M., 2007. Subjective expectations and asset return puzzles. Am. Econ. Rev. 97, 1102–1130.

Appendix 10.1: Solving the CCAPM with Growth[34]

Assume that there is a finite set of possible growth rates $\{g_1, \ldots, g_N\}$ whose realizations are governed by a Markov process with transition matrix **T** and entries π_{ij}. Then, for whatever g_i is realized in period $t+1$,

$$d_{t+1} = g_{t+1} Y_t = g_{t+1} c_t = g_i c_t$$

Under the usual utility specification, $U(c) = (c^{1-\gamma})/(1-\gamma)$, the basic equilibrium ($Y_t = c_t$) asset pricing equation reduces to

$$Y_t^{-\gamma} q(Y_t, g_i) = \delta \sum_{j=1}^{N} \pi_{ij} (g_j Y_t)^{-\gamma} [Y_t g_j + q(g_j Y_t, g_j)] \text{or}$$

$$q(Y_t, g_i) = \delta \sum_{j=1}^{N} \pi_{ij} \left(\frac{g_j Y_t}{Y_t}\right)^{-\gamma} \left[c_t g_j + q(g_j Y_t, g_j)\right]$$

Notice that the MRS is determined exclusively by the consumption growth rate.

The essential insight of Mehra and Prescott (1985) was to observe that a solution to this linear system has the form

$$q(Y_t, g_i) = v_i Y_t$$

for a set of constants $\{v_1, \ldots, v_N\}$, each identified with the corresponding growth rate.

With this functional form, the asset pricing equation reduces to

$$v_i Y_t = \delta \sum_{j=1}^{N} \pi_{ij} (g_j)^{-\gamma} [g_j Y_t + v_j g_j Y_t] \text{ or}$$

$$v_i = \delta \sum_{j=1}^{N} \pi_{ij} (g_j)^{1-\gamma} [1 + v_j] \tag{10.45}$$

This is again a system of linear equations in the N unknowns $\{v_1, \ldots, v_N\}$. Provided the growth rates are not too large (so that the agent's utility is not unbounded), a solution exists—a set of $\{v_1^*, \ldots, v_N^*\}$ that solves the system of Eqs. (10.45).

Thus, for any state $(Y, g_j) = (c, g_j)$, the equilibrium equity asset price is

$$q(Y, g_j) = v_j{}^* Y$$

If we suppose the current state is (Y, g_i) while next period it is $(g_j Y, g_j)$, then the one-period return earned by the equity security over this period is

[34] This appendix is based on Mehra and Prescott (1985).

$$r_{ij} = \frac{q(g_jY, g_j) + g_jY - q(Y, g_i)}{q(Y, g_i)}$$

$$= \frac{v_j^* g_j Y + g_j Y - v_i^* Y}{v_i^* Y}$$

$$= \frac{g_i(v_j^* + 1)}{v_i^*} - 1$$

and the mean or expected return, conditional on state i, is

$$r_i = \sum_{j=1}^{N} \pi_{ij} r_j$$

The unconditional equity return is thus given by

$$Er = \sum_{j=1}^{N} \hat{\pi}_j r_j$$

where $\hat{\pi}_j$ are the long-run stationary probabilities of each state.

The risk-free security is analogously priced as

$$q^{\text{rf}}(c, g_i) = \delta \sum_{j=1}^{N} \pi_{ij}(g_j)^{-\gamma}, \text{etc.}$$

Appendix 10.2: Some Properties of the Lognormal Distribution

Definition A10.1 A variable x is said to follow a lognormal distribution if $\ln x$ is normally distributed. Let $\ln x \sim N(\mu_x, \sigma_x^2)$. If this is the case,

$$E(x) = \exp\left\{\mu_x + \frac{1}{2}\sigma_x^2\right\}$$

$$E(x^a) = \exp\left\{a\mu_x + \frac{1}{2}a^2\sigma_x^2\right\}$$

$$\text{var}(x) = \exp\{2\mu_x + \sigma_x^2\}(\exp \sigma_x^2 - 1)$$

Suppose furthermore that x and y are two variables that are independently and identically lognormally distributed, then we also have

$$E(x^a y^b) = \exp\left\{a\mu_x + b\mu_y + \frac{1}{2}(a^2\sigma_x^2 + b^2\sigma_y^2) + 2\rho ab\sigma_x\sigma_y\right\}$$

where ρ is the correlation coefficient between $\ln x$ and $\ln y$.

Let us apply these relationships to consumption growth: g_t is lognormally distributed, i.e., $\ln g_t \sim N(\mu_g, \sigma_g^2)$.

We know that $E(g_t) = 1.0183$ and $\text{var}(g_t) = (0.0357)^2$. To identify (μ_g, σ_g^2), we need to find the solutions of

$$1.0183 = \exp\left\{\mu_g + \frac{1}{2}\sigma_g^2\right\}$$

$$(0.0357)^2 = \exp\{2\mu_g + \sigma_g^2\}(\exp\sigma_g^2 - 1)$$

Substituting the first equation squared into the second by virtue of the fact that $[\exp(y)]^2 = \exp(2y)]$ and solving for σ_g^2, one obtains

$$\sigma_g^2 = 0.00123$$

Substituting this value in the equation for μ_g, one solves for $\mu_g = 0.01752$.

We can directly use these values to solve Eq. (10.20):

$$E\{g_t^{-\gamma}\} = \exp\left\{-\gamma\mu_g + \frac{1}{2}\gamma^2\sigma_g^2\right\} = \exp\{-0.03258\} = 0.967945$$

thus $\delta = 1.024$.

Focusing now on the numerator of Eq. (10.21), one has

$$\exp\left\{\mu_g + \frac{1}{2}\sigma_g^2\right\}\exp\left\{-\gamma\mu_g + \frac{1}{2}\gamma^2\sigma_g^2\right\}$$

while the denominator is

$$\exp\left\{(1-\gamma)\mu_g + \frac{1}{2}(1-\gamma)^2\sigma_g^2\right\}$$

It remains to recall that $(\exp(a)\exp(b))/(\exp(c)) = \exp(a + b - c)$ to obtain Eq. (10.22).

Another application is as follows: The standard deviation of the pricing kernel $m_t = g_t^{-\gamma}$ where consumption growth g_t is lognormally distributed. Given that Em_t is as derived in Section 10.6, one estimates

$$\sigma^2(m_t) \cong \frac{1}{k}\left\{\sum_{i=1}^{k}[\delta(g_i)^{-\gamma} - Em_t]^2\right\}$$

for ln g_i drawn from $N(0.01752; 0.00123)$ and k sufficiently large (say $k = 10{,}000$). For $\gamma = 2$, one obtains $\sigma^2(m_t) = (0.00234)^2$, which yields

$$\frac{\sigma_m}{E\tilde{m}} \cong \frac{0.00234}{0.9559} = 0.00245$$