

**Ec2723, Asset Pricing I**  
**Class Notes, Fall 2005**

**Complete Markets, Incomplete Markets,  
and the Stochastic Discount Factor**

**John Y. Campbell<sup>1</sup>**

First draft: July 30, 2003  
This version: October 10, 2005

---

<sup>1</sup>Department of Economics, Littauer Center, Harvard University, Cambridge MA 02138, USA.  
Email [john\\_campbell@harvard.edu](mailto:john_campbell@harvard.edu).

## The SDF in a complete market

Consider a simple discrete-state model with states of nature  $s = 1 \dots S$ . We assume markets are complete, that is, for each state  $s$  a contingent claim is available that pays \$1 in state  $s$  and nothing in any other state. Write the price of this contingent claim as  $P_c(s)$ .

Any other asset is defined by its payoff  $X(s)$  in state  $s$ . We must have

$$P(X) = \sum_{s=1}^S P_c(s)X(s).$$

Multiply and divide by the probability of each state,  $\pi(s)$ :

$$P(X) = \sum_{s=1}^S \pi(s) \frac{P_c(s)}{\pi(s)} X(s) = \sum_{s=1}^S \pi(s) M(s) X(s) = \mathbb{E}[MX],$$

where  $M(s)$  is the ratio of state price to probability for state  $s$ , the *stochastic discount factor* or SDF. Any asset price is the expected product of the asset's payoff and the SDF.

Consider a riskless asset with payoff  $X(s) = 1$  in every state. The price

$$P_f = \sum_{s=1}^S P_c(s) = \mathbb{E}[M],$$

so the riskless interest rate

$$1 + R_f = \frac{1}{P_f} = \frac{1}{\mathbb{E}[M]}.$$

Now define *risk-neutral probabilities* or *pseudo-probabilities*

$$\pi^*(s) = (1 + R_f)P_c(s) = \frac{M(s)}{\mathbb{E}[M]}\pi(s).$$

We have  $\pi^*(s) > 0$  and  $\sum_s \pi^*(s) = 1$ , so they can be interpreted as if they were probabilities. We can rewrite the asset equation as

$$P(X) = \left( \frac{1}{1 + R_f} \right) \sum_{s=1}^S \pi^*(s) X(s) = \left( \frac{1}{1 + R_f} \right) \mathbb{E}^*[X].$$

The price of any asset is the pseudo-expectation of its payoff, discounted at the riskless interest rate.

### Utility maximization and the SDF

Consider an investor with initial wealth  $Y$  and income  $Y(s)$ . The investor's maximization problem is

$$\text{Max } u(C_0) + \sum_{s=1}^S \beta \pi(s) u(C(s))$$

subject to

$$C_0 + \sum_{s=1}^S P_c(s) C(s) = Y_0 + \sum_{s=1}^S P_c(s) Y(s).$$

Defining a Lagrange multiplier  $\lambda$  on the budget constraint, the first-order conditions are

$$\begin{aligned} u'(C_0) &= \lambda \\ \beta \pi(s) u'(C(s)) &= \lambda P_c(s) \text{ for } s = 1 \dots S. \end{aligned}$$

Thus

$$M(s) = \frac{P_c(s)}{\pi(s)} = \frac{\beta u'(C(s))}{u'(C_0)} = \frac{\beta u'(C(s))}{\lambda}$$

and

$$\frac{M(s_1)}{M(s_2)} = \frac{u'(C(s_1))}{u'(C(s_2))}.$$

The ratio of SDF realizations across states is the ratio of marginal utilities across states. Since this is true for any two investors  $i$  and  $j$ , we also have

$$\frac{u'_i(C_{t+1}^i)}{u'_i(C_t^i)} = \frac{u'_j(C_{t+1}^j)}{u'_j(C_t^j)},$$

assuming a common time discount factor  $\beta$  to keep the notation simple. This tells us that the consumption allocation in the economy is Pareto optimal, since a social planner allocating consumption to the two investors would solve

$$\text{Max } \lambda_i E \sum_t \beta^t u_i(C_t^i) + \lambda_j E \sum_t \beta^t u_j(C_t^j)$$

subject to  $C_t^i + C_t^j = C_t$ . The first-order condition

$$\lambda_i u'_i(C_t^i) = \lambda_j u'_j(C_t^j)$$

implies a common marginal utility ratio across states. This condition characterizes perfect risk-sharing in a complete markets economy. Note that it holds ex post, not just ex ante, and is therefore a very strong statement about individual agents' consumption.

### **The idea of the “martingale method”**

The above logic has been applied to solve portfolio choice problems. In a model with only financial wealth and a single period,

$$C_{t+1}^j = X_{t+1}^j,$$

where  $X_{t+1}^j$  is the payoff on investor  $j$ 's portfolio. Given complete markets there is a unique SDF  $M_{t+1}$  such that

$$M_{t+1} = \frac{\beta}{\lambda_j} u'_j(X_{t+1}^j) \implies X_{t+1}^j = u_j'^{-1} \left( \frac{\lambda_j}{\beta} M_{t+1} \right).$$

We solve for the  $\lambda_j$  that makes the payoff  $X_{t+1}^j$  affordable at time  $t$  given the investor's current wealth. Then the investor holds a portfolio of contingent claims that delivers  $X_{t+1}^j$  at time  $t + 1$ . This method has been developed in continuous time by Cox and Huang (1989) and is explained in Campbell and Viceira (2002), Chapter 5.

### **Existence of a representative agent with complete markets**

With complete markets, the condition

$$\frac{u'_i(C_{t+1}^i)}{u'_i(C_t^i)} = \frac{u'_j(C_{t+1}^j)}{u'_j(C_t^j)}$$

ensures that all agents have the same ordering of marginal utility across states. With declining marginal utility, this means that they all have the same ordering of consumption across states. Without loss of generality, number states such that

$$C^i(s_1) \leq C^i(s_2) \leq \dots \leq C^i(s_S)$$

for all agents  $i$ . Define aggregate consumption  $C(s) = \sum_i C^i(s)$ . Then we have

$$C(s_1) \leq C(s_2) \leq \dots \leq C(s_S).$$

Also, we have

$$M(s_1) \geq M(s_2) \geq \dots \geq M(s_S).$$

Now find a function  $g(C(s))$  s.t.

$$\frac{g(C(s_j))}{g(C(s_k))} = \frac{M(s_j)}{M(s_k)}$$

for all states  $j$  and  $k$ . The above ordering conditions ensure that this is always possible, with  $g > 0$  and  $g' \leq 0$ . Finally, integrate to find a function  $v(C(s))$  s.t.

$$v'(C(s)) = g(C(s)).$$

The function  $v(\cdot)$  is the utility function of a representative agent who consumes aggregate consumption and holds the market portfolio of all wealth.

This argument also shows that with complete markets, the market portfolio is *efficient* (a generalization of the concept of mean-variance efficiency). That is, there exists some concave utility function that would induce an agent to hold the market portfolio. This is *not* true in general with incomplete markets.

Unfortunately the above construction does not tell us anything about the form of the representative agent utility function. In general it need not resemble the utility functions of individual agents. Only in tractable special cases (e.g. CARA with normal risks) do we get a representative-agent utility function in the same class as the utility functions of individual agents.

## The SDF in incomplete markets

What if markets are incomplete? We continue to observe a set of payoffs  $X$  and prices  $P$ . The set of all payoffs (the payoff space) is  $\Xi$ . We assume:

(A1) Portfolio formation  $X_1, X_2 \in \Xi \implies aX_1 + bX_2 \in \Xi$  for any real  $a, b$ .

(A2) Law of One Price  $P(aX_1 + bX_2) = aP(X_1) + bP(X_2)$ .

*Theorem.*  $A1, A2 \implies$  there exists a unique payoff  $X^* \in \Xi$  s.t.  $P(X) = E(X^*X)$  for all  $X \in \Xi$ .

Sketch of proof: Assume that there are  $N$  basis payoffs  $X_1, \dots, X_N$ . Construct a vector  $X = [X_1 \dots X_N]'$ . Write the set  $\Xi = \{c'X\}$ . We want to find some vector  $X^* = c'X$  that prices the basis payoffs. That is, we want

$$P = E[X^*X] = E[XX'c]$$

which requires

$$c = E[XX']^{-1}P$$

and

$$X^* = P'E[XX']^{-1}X.$$

This construction for  $X^*$  always exists and unique provided that the matrix  $E[XX']$  is nonsingular.

Notes:

- We can subtract means and rewrite all of this in terms of covariance matrices.
- Only the SDF that is a linear combination of asset payoffs is unique. There may be many other SDF's of the form  $M = X^* + \epsilon$ , where  $E[\epsilon X] = 0$ . These must all have higher variance than  $X^*$  (Hansen-Jagannathan variance bound).
- $X^*$  is the projection of every SDF onto the space of payoffs. Thus it can be thought of as the portfolio of assets that best mimics the behavior of every SDF. (For more on this, see the discussion of the benchmark portfolio below.)

*Definition.* A payoff space  $\Xi$  and pricing function  $P(X)$  have absence of arbitrage if all  $X$  s.t.  $X \geq 0$  always and s.t.  $X > 0$  with positive probability have  $P(X) > 0$ .

*Theorem.*  $P = E(MX)$  and  $M(s) > 0 \implies$  absence of arbitrage.

Proof:  $P(X) = \sum_s \pi(s)M(s)X(s)$ , and no terms in this expression are ever negative.

*Theorem.* Absence of arbitrage  $\implies \exists M$  s.t.  $P = E(MX)$  and  $M(s) > 0$ .

Proof: See Cochrane, *Asset Pricing*, Chapter 4, for a geometric proof. The intuition is that with absence of arbitrage, we can always find a complete-markets, contingent-claims economy (in general, many such economies) that could have generated the asset prices we observe.

### The SDF and risk premia

For a general risky asset  $i$ , we have

$$\begin{aligned} P_{it} &= E_t[M_{t+1}X_{i,t+1}] = E_t[M_{t+1}]E_t[X_{i,t+1}] + \text{Cov}_t(M_{t+1}, X_{i,t+1}) \\ &= \frac{E_t[X_{i,t+1}]}{(1 + R_{f,t+1})} + \text{Cov}_t(M_{t+1}, X_{i,t+1}). \end{aligned}$$

For assets with positive prices, we can divide through by  $P_{it}$  and use  $(1 + R_{i,t+1}) = X_{i,t+1}/P_{i,t+1}$  to get

$$\begin{aligned} 1 &= E_t[M_{t+1}(1 + R_{i,t+1})] = E_t[M_{t+1}]E_t[1 + R_{i,t+1}] + \text{Cov}_t(M_{t+1}, R_{i,t+1}) \\ E_t[1 + R_{i,t+1}] &= (1 + R_{f,t+1})(1 - \text{Cov}_t(M_{t+1}, R_{i,t+1})). \\ E_t(R_{i,t+1} - R_{f,t+1}) &= \frac{-\text{Cov}_t(M_{t+1}, R_{i,t+1} - R_{f,t+1})}{E_t M_{t+1}}. \end{aligned}$$

Sometimes it is convenient to work with log versions of these equations, assuming joint lognormality of asset returns and the SDF. Taking logs, the riskless interest rate

$$r_{f,t+1} = -E_t m_{t+1} - \sigma_{mt}^2/2,$$

where  $r_{f,t+1} \equiv \log(1 + R_{f,t+1})$ ,  $m_{t+1} \equiv \log(M_{t+1})$ , and  $\sigma_{mt}^2 = \text{Var}_t(m_{t+1})$ . Similarly the log risk premium, adjusted for Jensen's Inequality by adding one-half the own variance, is

$$E_t r_{i,t+1} - r_{f,t+1} + \sigma_i^2/2 = -\sigma_{imt},$$

where  $\sigma_{imt} \equiv \text{Cov}_t(r_{i,t+1}, m_{t+1})$ . In intertemporal equilibrium models such as the representative-agent model with power utility these equations are often easier to work with.

## Volatility bounds on the SDF

We can use the fact that the correlation between the SDF and any excess return must be greater than minus one to obtain a lower bound on the volatility of the SDF. We have

$$E_t(R_{i,t+1} - R_{f,t+1}) = \frac{-\text{Cov}_t(M_{t+1}, R_{i,t+1} - R_{f,t+1})}{E_t M_{t+1}} \leq \frac{\sigma_t(M_{t+1})\sigma_t(R_{i,t+1} - R_{f,t+1})}{E_t M_{t+1}}.$$

Rearranging, we get

$$\frac{\sigma_t(M_{t+1})}{E_t M_{t+1}} \geq \frac{E_t(R_{i,t+1} - R_{f,t+1})}{\sigma_t(R_{i,t+1} - R_{f,t+1})}.$$

The Sharpe ratio for asset  $i$  puts a lower bound on the volatility of the stochastic discount factor. The tightest lower bound is achieved by finding the risky asset, or portfolio of assets, with the highest Sharpe ratio. This bound was first stated by Shiller (1982).

The logarithmic version of this bound is even easier. Since  $-\sigma_{imt} \leq \sigma_{it}\sigma_{mt}$ ,

$$\sigma_{mt} \geq \frac{E_t r_{i,t+1} - r_{f,t+1} + \sigma_i^2/2}{\sigma_{it}}.$$

The ratio on the RHS is in the range 0.33 to 0.5 for the aggregate US stock market in the 20th Century, implying a large standard deviation for the SDF. Recall that the mean SDF must be close to 1 and the SDF is always positive, so a volatility of 0.33 to 0.5 implies that the lower bound is only 2 or 3 standard deviations below the mean. It is surprising that marginal utility routinely changes this much in a year. This is the most general way to understand the famous *equity premium puzzle*.

Hansen and Jagannathan (1991) extended this idea and related it to mean-variance analysis. Suppose there are  $N$  risky assets and no riskless asset, so the mean of the SDF is not pinned down by the mean return on any asset. Write this unknown mean SDF as  $\bar{M}$ . Start from

$$\iota = E[(\iota + R_t)M_t]$$

and write  $\Sigma$  for the variance-covariance matrix of asset returns. The minimum-variance stochastic discount factor is a linear combination of asset returns:

$$M_t^*(\bar{M}) = \bar{M} + (R_t - R)' \beta(\bar{M})$$

for some coefficient vector  $\beta(\bar{M})$ . We must have

$$\iota = E[(\iota + R_t)M_t^*(\bar{M})] = (\iota + \bar{R})\bar{M} + \text{Cov}(R_t, M_t^*(\bar{M})) = (\iota + \bar{R})\bar{M} + \Sigma\beta(\bar{M}).$$

Rearranging,

$$\beta(\bar{M}) = \Sigma^{-1}(\iota - (\iota + \bar{R})\bar{M})$$

and

$$\begin{aligned} \text{Var}(M_t^*(\bar{M})) &= \beta(\bar{M})'\Sigma\beta(\bar{M}) = (\iota - (\iota + \bar{R})\bar{M})'\Sigma^{-1}(\iota - (\iota + \bar{R})\bar{M}) \\ &= \bar{M}^2(\iota + \bar{R})'\Sigma^{-1}(\iota + \bar{R}) - 2\bar{M}\iota'\Sigma^{-1}(\iota + \bar{R}) + \iota'\Sigma^{-1}\iota \\ &= A\bar{M}^2 - 2B\bar{M} + C, \end{aligned}$$

where  $A = (\iota + \bar{R})'\Sigma^{-1}(\iota + \bar{R})$ ,  $B = \iota'\Sigma^{-1}(\iota + \bar{R})$ , and  $C = \iota'\Sigma^{-1}\iota$  are just as we defined them in the standard mean-variance analysis, except that here we are working with gross mean returns. (We could have defined them that way in the mean-variance analysis also, and obtained the same formulas.)

The variance of  $M_t^*(\bar{M})$  gives a lower bound on the variance of any SDF  $M_t(\bar{M})$  with the same mean. Any such SDF must satisfy

$$E[(\iota + R_t)(M_t(\bar{M}) - M_t^*(\bar{M}))] = \text{Cov}(R_t, M_t(\bar{M}) - M_t^*(\bar{M})) = 0.$$

Since  $M_t^*(\bar{M})$  is itself a linear combination of returns, this implies that

$$\text{Cov}(M_t^*(\bar{M}), M_t(\bar{M}) - M_t^*(\bar{M})) = 0.$$

Then

$$\text{Var}(M_t(\bar{M})) = \text{Var}(M_t^*(\bar{M})) + \text{Var}(M_t(\bar{M}) - M_t^*(\bar{M})) \geq \text{Var}(M_t^*(\bar{M})).$$

If we augment the set of risky asset returns with a hypothetical riskless return  $1/\bar{M}$ , then we can define a benchmark return

$$1 + R_{bt}(\bar{M}) = \frac{M_t^*(\bar{M})}{E[M_t^*(\bar{M})^2]}.$$

The benchmark return has the following properties:

- It lies on the minimum-variance frontier (the lower part, not the mean-variance efficient frontier).

- It has the highest possible correlation with the SDF.
- Beta pricing works with the benchmark return.
- 

$$\frac{1/\bar{M} - (1 + \bar{R}_b)}{\sigma_b} \leq \frac{\sigma_M(\bar{M})}{\bar{M}}.$$

This allows us to link the geometry of the mean-variance frontier and the geometry of the Hansen-Jagannathan frontier in an elegant way.

### Factor structure of the SDF

The SDF approach is another way to understand multifactor models. Assume that the SDF is a linear combination of  $K$  common factors  $f_{k,t+1}$ ,  $k = 1 \dots K$ . For simplicity assume that the factors have conditional mean zero and are orthogonal to one another. If

$$M_{t+1} = a_t - \sum_{k=1}^K b_{kt} f_{k,t+1},$$

then

$$-\text{Cov}_t(M_{t+1}, R_{i,t+1} - R_{f,t+1}) = \sum_{k=1}^K b_{kt} \sigma_{ikt} = \sum_{k=1}^K (b_{kt} \sigma_{kt}^2) \left( \frac{\sigma_{ikt}}{\sigma_{kt}^2} \right) = \sum_{k=1}^K \lambda_{kt} \beta_{ikt}.$$

Here  $\sigma_{ikt}$  is the conditional covariance of asset return  $i$  with the  $k$ 'th factor,  $\sigma_{kt}^2$  is the conditional variance of the  $k$ 'th factor,  $\lambda_{kt} \equiv b_{kt} \sigma_{kt}^2$  is the price of risk of the  $k$ 'th factor, and  $\beta_{ikt} \equiv \sigma_{ikt} / \sigma_{kt}^2$  is the beta or regression coefficient of asset return  $i$  on that factor.

Note how this is consistent with earlier insights about multifactor models:

- Single-period model with quadratic utility implies consumption equals wealth and marginal utility is linear. Thus the SDF must be linear in future wealth, or equivalently the market portfolio return.
- In a single-period model with  $K$  common shocks and completely diversifiable idiosyncratic risk, marginal utility and hence the SDF can depend only on the common shocks.

Note the importance of conditioning information. A conditional multifactor model does not generally imply an unconditional multifactor model of the same form. The relevant covariance for an unconditional model is the unconditional covariance  $-\text{Cov}(M_{t+1}, R_{i,t+1} - R_{f,t+1}) = -\text{Cov}(a_t - \sum_{k=1}^K b_{kt} f_{k,t+1}, R_{i,t+1} - R_{f,t+1})$ , and this involves covariances of the coefficients  $a_t$  and  $b_t$  with returns as well as covariances of the factors  $f_{k,t+1}$  with returns. One way to handle this problem is to model the coefficients themselves as linear functions of observable instruments:  $a_t = a'z_t$  and  $b_t = b'z_t$ , where  $z_t$  is a vector of instruments including a constant. In this case one obtains an unconditional multifactor model in which the factors include the original  $f_{k,t+1}$ , the instruments  $z_t$ , and all cross-products of  $f_{k,t+1}$  and  $z_t$ . See Cochrane, Chapter 8 for details.

### **Time series models of the SDF**

Predictability of asset returns requires predictability in the SDF. To get time-series variation in the riskless real interest rate, we need a changing conditional mean of the SDF. To get time-series variation in risk premia, we need changing covariances of asset returns with the SDF. To get time-series variation in the risk premium on an asset that is perfectly correlated with the SDF and has a constant return variance, we need time-variation in the variance of the SDF.

Thus dynamic asset pricing models often involve modelling both first and second conditional moments of the SDF, drawing on both conventional linear time-series analysis and nonlinear models of conditional volatility such as ARCH models.

## Generalized Method of Moments

Hansen's Generalized Method of Moments (GMM) is an econometric approach that is particularly well suited for testing models of the SDF. If the model says that

$$P_t = E_t[M_{t+1}(b)X_{t+1}],$$

where  $b$  is a vector of parameters, then we can take unconditional expectations to get

$$E[P_t] = E[M_{t+1}(b)X_{t+1}]$$

or

$$E[M_{t+1}(b)X_{t+1} - P_t] = E[u_{t+1}(b)] = 0.$$

Here  $u_{t+1}(b)$  is the pricing error of the model, which depends on the parameters of the model for the SDF. If we have a vector of asset returns, then  $u_{t+1}(b)$  is a vector.

There are several variants of this idea. For example, we might divide through by prices, expressing the model in terms of returns:

$$E[M_{t+1}(b)(1 + R_{t+1}) - 1] = 0.$$

Or we might premultiply the conditional equation by instruments  $Z_t$  known at time  $t$ , then take unconditional expectations to get

$$E[Z_t M_{t+1}(b)(1 + R_{t+1}) - Z_t] = 0.$$

In all these cases we will write  $u_{t+1}(b)$  for the pricing error, the object that should have unconditional expectation equal to zero if the SDF model is correct. With  $N$  asset returns and  $J$  instruments,  $u_{t+1}(b)$  is an  $NJ \times 1$  vector.

We define  $g_T(b)$  as the sample mean of the  $u_{t+1}(b)$  in a sample of size  $T$ :

$$g_T(b) = \frac{1}{T} \sum_{t=1}^T u_{t+1}(b).$$

The *first-stage GMM estimate* of  $b$  solves

$$\min g_T(b)'Wg_T(b)$$

for some weighting matrix  $W$ . This estimate,  $\hat{b}_1$ , is consistent and asymptotically normal.

We can continue to a second stage. Using  $\hat{b}_1$  we form an estimate  $\hat{S}$  of

$$S = \sum_{j=-\infty}^{\infty} \text{E}[u_{t+1}(b)u_{t+1-j}(b)'].$$

Then the *second-stage GMM estimate* of  $b$  solves

$$\min g_T(b)'\hat{S}^{-1}g_T(b).$$

This estimate,  $\hat{b}_2$ , is consistent, asymptotically normal, and asymptotically efficient (that is, it has the smallest variance-covariance matrix among all choices of weighting matrix  $W$ ).

The variance-covariance matrix of  $\hat{b}_2$  is given by

$$\text{Var}(\hat{b}_2) = \frac{1}{T}(d'S^{-1}d)^{-1},$$

where

$$d = \frac{\partial g_T(b)}{\partial b},$$

and we can estimate it consistently using sample estimates  $\hat{S}^{-1}$  and  $\hat{d}$ . Also, the model can be tested using the following result for the distribution of the minimized second-stage objective function:

$$T \min g_T(b)'\hat{S}^{-1}g_T(b) \sim \chi^2(NJ - K),$$

where  $NJ$  is the number of moment conditions and  $K$  is the number of parameters.

This framework is extremely general. The usual formulas for OLS and GLS regression can be derived as special cases. The decision whether to proceed to a second-stage GMM estimate is equivalent to the decision whether to use GLS. The second-stage estimate is more efficient asymptotically, but can behave poorly in finite samples if the matrix  $S$  is poorly estimated. This would be the case, for example, if the number of moment conditions is large relative to the sample size.

Formulas for the variance-covariance matrix of  $\hat{b}_1$  and for the distribution of the minimized first-stage objective function are given in Cochrane, Chapter 11. These make it possible to do valid statistical inference using first-stage GMM estimates (analogous to correcting the standard errors of OLS regression for the presence of correlated errors).

Cochrane, Chapters 12 and 13 show how traditional regression-based tests of linear factor models can be understood within the GMM framework. Cochrane, Chapters 14 and 15 compare these approaches with maximum likelihood. The general theme is that first-stage GMM estimates with a sensible weighting matrix are often more persuasive than more sophisticated two-stage GMM or maximum likelihood estimates.

What is a sensible weighting matrix to use in the first stage? One choice is the identity matrix. Another, advocated by Hansen and Jagannathan (*Journal of Finance*, 1997) is the inverse of the variance-covariance matrix of asset returns. This has some of the advantages of the optimal weighting matrix but can be calculated without reference to first-stage parameter estimates.