

Chapter 3 Asset Pricing Theories, Models, and Tests

NIKOLAY GOSPODINOV

Associate Professor, Concordia University and CIREQ

CESARE ROBOTTI

Financial Economist and Associate Policy Adviser, Federal Reserve Bank of Atlanta
Affiliate Professor, EDHEC Risk Institute

KEYWORDS: Stochastic discount factor; two-pass cross-sectional regressions; asset pricing models; misspecified models; useless factors.

ABSTRACT

An important but still partially unanswered question in the investment field is why different assets earn substantially different returns on average. Financial economists have typically addressed this question in the context of theoretically or empirically motivated asset pricing models. Since many of the proposed “risk” theories are plausible, a common practice in the literature is to take the models to the data and perform “horse races” among competing asset pricing specifications. A “good” asset pricing model should produce small pricing (expected return) errors on a set of test assets and should deliver reasonable estimates of the underlying market and economic risk premia. This chapter provides an up-to-date review of the statistical methods that are typically used to estimate, evaluate, and compare competing asset pricing models. The analysis also highlights several pitfalls in the current econometric practice and offers suggestions for improving empirical tests.

INTRODUCTION

Many asset pricing theories predict that the price of an asset should be lower (its expected return higher) if the asset provides a poor hedge against changes in future market conditions (Rubinstein, 1976; Breeden, 1979). The classic capital asset pricing model (CAPM) of Sharpe (1964) and Lintner (1965) considers the case in which investment opportunities are constant

and investors hold efficient portfolios so as to maximize their expected return for a given level of variance. The CAPM predicts that an asset's risk premium will be proportional to its beta – the measure of return sensitivity to the aggregate market portfolio return. The considerable empirical evidence against the CAPM points to the fact that variables other than the rate of return on a market-portfolio proxy command significant risk premia. The theory of the intertemporal CAPM (ICAPM) (Merton, 1973; Long, 1974) suggests that these additional variables should proxy for the position of the investment opportunity set. Although the ICAPM does not identify the various state variables, leading Fama (1991) to label the ICAPM as a “fishing license,” Breeden (1979) shows that Merton's ICAPM is actually equivalent to a single-beta consumption model (CCAPM) since the chosen level of consumption endogenously reflects the various hedging-demand effects of the ICAPM.

Over the years, researchers have made many attempts to refine the theoretical predictions and improve the empirical performance of the CAPM and CCAPM. Popular extensions include internal and external habit models (Abel, 1990; Constantinides, 1990; Ferson and Constantinides, 1991; Campbell and Cochrane, 1999), models with non-standard preferences and rich consumption dynamics (Epstein and Zin, 1989, 1991; Weil, 1989; Bansal and Yaron, 2004), models that allow for slow adjustment of consumption to the information driving asset returns (Parker and Julliard, 2005), conditional models (Jagannathan and Wang, 1996; Lettau and Ludvigson, 2001), disaster risk models (Berkman, Jacobsen, and Lee, 2011), and the well-known “three-factor model” of Fama and French (1993). Although empirical observation primarily motivated the Fama-French model, its size and book-to-market factors are sometimes viewed as proxies for more fundamental economic variables.

The asset pricing theories listed above, to be of practical interest, need to be confronted with the data. Two main econometric methodologies have emerged to estimate and test asset pricing models: (1) the generalized method of moments (GMM) methodology for models written

in stochastic discount factor (SDF) form and (2) the two-pass cross-sectional regression (CSR) methodology for models written in beta form.

The SDF approach to asset pricing indicates that the price of a security is obtained by "discounting" its future payoff by a valid SDF so that the expected present value of the payoff is equal to the current price. In practice, finding a valid SDF, i.e., an SDF that prices each asset correctly, is impossible and researchers have to rely on some candidate SDFs to infer the price of an asset. Although testing whether a particular asset pricing model is literally true is interesting, a more useful task for empirical researchers is to determine how wrong a model is and to compare the performance of competing asset pricing models. The latter task requires a scalar measure of model misspecification. While many reasonable measures can be used, the one introduced by Hansen and Jagannathan (1997) has gained tremendous popularity in the empirical asset pricing literature. Many researchers have used their proposed measure, called the Hansen-Jagannathan distance (HJ-distance), both as a model diagnostic and as a tool for model selection. Examples include Jagannathan and Wang (1996), Jagannathan, Kubota, and Takehara (1998), Campbell and Cochrane (2000), Lettau and Ludvigson (2001), Hodrick and Zhang (2001), Dittmar (2002), Farnsworth, Ferson, Jackson, and Todd (2002), Chen and Ludvigson (2009), Kan and Robotti (2009), Li, Xu, and Zhang (2010), and Gospodinov, Kan, and Robotti (2011a). Asset pricing models in SDF form are generally estimated and tested using GMM methods. Importantly, the SDF approach and the HJ-distance metric are applicable whether or not the pricing model is linear in a set of systematic risk factors.

When a model specifies that asset expected returns are linear in the betas (beta-pricing model), the CSR method proposed by Black, Jensen, and Scholes (1972) and Fama and MacBeth (1973) has been the preferred method in empirical finance given its simplicity and intuitive appeal. Although there are many variations of the CSR methodology, the basic approach always involves two steps or passes. In the first pass, the betas of the test assets are estimated using the usual ordinary least squares (OLS) time series regression of returns on

some common factors. In the second pass, the returns on the test assets are regressed on the betas estimated from the first pass. Running this second-pass CSR on a period-by-period basis enables obtaining the time series of the intercept and the slope coefficients. The average values of the intercept and the slope coefficients are then used as estimates of the zero-beta rate (expected return for risky assets with no systematic risk) and factor risk premia, with standard errors computed from these time series as well. Given its simple intuitive appeal, the most popular measure of model misspecification in the CSR framework has been the R^2 for the cross-sectional relation (Kandel and Stambaugh, 1995; Kan, Robotti, and Shanken, 2010). This R^2 indicates the extent to which the model's betas account for the cross-sectional variation in average returns, typically for a set of asset portfolios.

After reviewing the SDF and beta approaches to asset pricing, this chapter describes several pitfalls in the current econometric analyses and provides suggestions for improving empirical tests. Particular emphasis is given to the role played by model misspecification and to the need for more reliable inference procedures in estimating and evaluating asset pricing models.

STOCHASTIC DISCOUNT FACTOR REPRESENTATION

The SDF approach to asset pricing provides a unifying framework for pricing stocks, bonds, and derivative products and is based on the following fundamental pricing equation (Cochrane, 2005):

$$p_t = E_t[m_{t+1}x_{t+1}], \quad (3.1)$$

where p_t is an N -vector of asset prices at time t ; $x_{t+1} = p_{t+1} + d_{t+1}$ is an N -vector of asset payoffs with d_{t+1} denoting any asset's dividend, interest or other payment received at time $t + 1$; m_{t+1} is an SDF, which depends on data and parameters; and E_t is a conditional expectation given all publicly available information at time t .

Dividing both sides of the fundamental pricing equation by p_t (assuming non-zero prices) and rearranging yields

$$E_t[m_{t+1}(1 + R_{t+1}) - 1_N] = 0_N, \quad (3.2)$$

where $R_{t+1} = \frac{x_{t+1}}{p_t} - 1 = \frac{p_{t+1} + d_{t+1}}{p_t} - 1$ is an N -vector of asset returns and 1_N and 0_N are N -vectors of ones and zeros, respectively.

Portfolios based on excess returns, $R_{t+1}^e = R_{t+1} - R_t^f 1_N$, where R_t^f denotes the risk-free rate at time t , are called zero-cost portfolios. Since the risk-free rate is known ahead of time, it follows that $E_t[m_{t+1}(1 + R_t^f)] = E_t[m_{t+1}](1 + R_t^f) = 1$ and $E_t[m_{t+1}] = \frac{1}{(1 + R_t^f)}$. In this case, with zero prices and payoffs R_{t+1}^e , the fundamental pricing equation is given by

$$E_t[m_{t+1} R_{t+1}^e] = 0_N. \quad (3.3)$$

As an example of the SDF approach, consider the problem of a representative agent maximizing her lifetime expected utility

$$\sum_{t=1}^{\infty} \beta^t E_0[u(c_t)] \quad (3.4)$$

subject to a budget constraint

$$a_{t+1} = (a_t + y_t - c_t)(1 + R_{t+1}), \quad (3.5)$$

where β , c_t , a_t and y_t denote the time preference parameter, consumption, asset's amount and income at time t , respectively. The first-order condition for the optimal consumption and portfolio choice is given by

$$E_t \left[\beta \frac{u'(c_{t+1})}{u'(c_t)} (1 + R_{t+1}) - 1_N \right] = 0_N, \quad (3.6)$$

where $u'(c)$ denotes the first derivative of the utility function $u(c)$ with respect to c . This first-order condition takes the form of the fundamental pricing equation with SDF given by the intertemporal marginal rate of substitution

$$m_{t+1} = \beta \frac{u'(c_{t+1})}{u'(c_t)}. \quad (3.7)$$

While the SDF in Equation 3.7 is positive by construction, an SDF can possibly price assets correctly and, at the same time, take on negative values, especially when the SDF is linear in a set of risk factors. Although a negative SDF does not necessarily imply the existence of arbitrage opportunities, dealing with positive SDFs is generally desirable, especially when interest lies in pricing derivatives (positive payoffs should have positive prices). Therefore, a common practice in the derivative pricing literature is to consider Equation 3.1 with $m_{t+1} > 0$, which implies the absence of arbitrage. In some situations, however, imposing this positivity constraint can be problematic. For example, if one is interested in comparing the performance of competing asset pricing models on a given set of test assets using the distance metric proposed by Hansen and Jagannathan (1997), constraining the admissible SDF to be positive is not very meaningful. Gospodinov, Kan, and Robotti (2010a) provide a rigorous analysis of the merits and drawbacks of the no-arbitrage HJ-distance metric.

BETA REPRESENTATION

By the law of iterated expectations, the conditional form of the fundamental pricing equation for gross returns can be reduced to its unconditional counterpart:

$$E[m_{t+1}(1 + R_{t+1})] = 1_N. \quad (3.8)$$

From the covariance decomposition (suppressing the time index for simplicity), the pricing equation for asset i can be rewritten as

$$1 = E[m(1 + R^i)] = E[m]E[1 + R^i] + Cov[m, (1 + R^i)]. \quad (3.9)$$

Then, dividing both sides by $E[m] > 0$ and rearranging,

$$E[R^i] = \frac{1}{E[m]} + \frac{Cov[m, R^i]}{Var[m]} \left[-\frac{Var[m]}{E[m]} \right] = \gamma_0 + \beta_{i,m} \lambda_m, \quad (3.10)$$

using that $\frac{1}{E[m]} = 1 + R^f = 1 + \gamma_0$ from above. Note that $\beta_{i,m} = \frac{Cov[m, R^i]}{Var[m]}$ is the regression coefficient of the return R^i on m and $\lambda_m = -\frac{Var[m]}{E[m]} < 0$ denotes the price of risk.

Recall that the SDF m is a function of the data and parameters. Suppose now that m can be approximated by a linear function of K (risk) factors f that serve as proxies for marginal utility growth:

$$m = \tilde{f}'\theta, \quad (3.11)$$

where $\tilde{f} = (1, f')'$. Then, substituting for m into the fundamental pricing equation and rearranging (see Cochrane, 2005, pp.107–108),

$$E[R^i] = \gamma_0 + \gamma_1'\beta_i, \quad (3.12)$$

where the β_i 's are the multiple regression coefficients of R^i on f and a constant, γ_0 is the zero-beta rate and γ_1 is the vector of risk premia on the K factors. The *beta* representation of a factor pricing model can be rewritten in compact form as

$$E[R] = B\gamma, \quad (3.13)$$

where $B = [1_N, \beta]$, $\beta = Cov[R, f]Var[f]^{-1}$ is an $(N \times K)$ matrix of factor loadings and $\gamma = (\gamma_0, \gamma_1')'$. Constant portfolio characteristics can easily be accommodated in Equation 3.13 (Kan et al., 2010). Jagannathan, Skoulakis, and Wang (2010) show how to write the beta-pricing relation when characteristics are time-varying.

For ease of exposition, the following analysis will mostly focus on the case of linear asset pricing models, but the techniques in this chapter are applicable to nonlinear models as well.

GMM ESTIMATION AND EVALUATION OF ASSET PRICING MODELS IN SDF FORM

Using Equation 3.11, the pricing errors of the N test assets can be expressed as

$$g(\theta) = E[m(1 + R)] - 1_N = E[(1 + R)\tilde{f}'\theta] - 1_N = D\theta - 1_N, \quad (3.14)$$

where $D = E[(1 + R)\tilde{f}']$. Let $t = 1, 2, \dots, T$ denote the number of time series observations on the test assets and the factors. The sample analog of the pricing errors is given by

$$g_T(\theta) = \frac{1}{T} \sum_{t=1}^T (1 + R_t)\tilde{f}_t'\theta - 1_N. \quad (3.15)$$

For a given weighting matrix W_T , the GMM estimator of θ minimizes the quadratic form

$$g_T(\theta)'W_T g_T(\theta) \quad (3.16)$$

and solves the first-order condition

$$D_T'W_T(D_T\theta - 1_N) = 0, \quad (3.17)$$

where $D_T = \frac{\partial g_T(\theta)}{\partial \theta'} = \frac{1}{T} \sum_{t=1}^T (1 + R_t) \tilde{f}_t'$. Solving this system of linear equations for θ yields

$$\hat{\theta} = (D_T'W_T D_T)^{-1}(D_T'W_T 1_N). \quad (3.18)$$

The optimal GMM estimator (under the assumption that the model is correctly specified) sets $W_T = V_T^{-1}$, where $V_T = \text{Var}[T^{-1/2}g_T(\theta)]$. In this case,

$$\hat{\theta} = (D_T'V_T^{-1}D_T)^{-1}(D_T'V_T^{-1}1_N), \quad (3.19)$$

where V_T is evaluated at some preliminary (consistent) estimator $\tilde{\theta}$ (typically obtained using $W_T = I_N$). If the model is correctly specified, i.e., it explains the test assets correctly, the pricing errors $g(\theta) = E[(1 + R)\tilde{f}'\theta] - 1_N$ are zero and the model's restrictions can be tested using the statistic

$$Tg_T(\hat{\theta})'V_T^{-1}g_T(\hat{\theta}) \rightarrow^d \chi_{(N-K-1)}^2. \quad (3.20)$$

If the model is misspecified, the value of the test statistic depends on the choice of W_T . Therefore, for model comparison, using the same W_T across models makes more sense.

Hansen and Jagannathan (1997) suggest using $W = U^{-1}$, where $U = E[(1 + R)(1 + R)']$ is the second moment matrix of the gross returns with sample analog U_T . Then, the sample HJ-distance is defined as

$$\delta_T(\theta) = \sqrt{g_T(\theta)'U_T^{-1}g_T(\theta)}. \quad (3.21)$$

and

$$\hat{\theta} = \text{argmin}_{\theta \in \Theta} \delta_T(\theta) = (D_T'U_T^{-1}D_T)^{-1}(D_T'U_T^{-1}1_N) \quad (3.22)$$

is the resulting GMM estimator. The HJ-distance has an interesting economic interpretation: (1) it measures the minimum distance between the proposed SDF and the set of valid SDFs, and (2) it represents the maximum pricing error of a portfolio of returns with unit second moment.

The HJ-distance test of correct model specification is based on (Jagannathan and Wang, 1996; Parker and Julliard, 2005)

$$T\delta_T^2(\hat{\theta}) \rightarrow^d \sum_{j=1}^{N-K-1} \xi_j v_j, \quad (3.23)$$

where the v_j 's are independent χ_1^2 random variables and the ξ_j 's are the non-zero eigenvalues of the matrix

$$V^{1/2}U^{-1/2} \left[I_N - (U^{-1/2})' D(D'U^{-1}D)^{-1} D'U^{-1/2} \right] (U^{-1/2})' (V^{1/2})'. \quad (3.24)$$

If $T\delta_T^2(\hat{\theta})$ exceeds the critical value from this weighted chi-squared distribution, then the model is misspecified. In this case, the traditional standard errors of the estimates $\hat{\theta}$ proposed by Hansen (1982) need to be adjusted for model misspecification (Kan and Robotti, 2009; Gospodinov et al., 2011a). Even if all candidate asset pricing models are misspecified, knowing which model provides the smaller pricing errors is still interesting. The statistical comparison of the HJ-distances of two or more competing models depends on whether the models are correctly specified or misspecified, nested or non-nested. Kan and Robotti (2009) and Gospodinov et al. (2011a) provide model selection tests to compare the performance of linear and nonlinear asset pricing models.

BETA-PRICING MODELS AND TWO-PASS CROSS-SECTIONAL REGRESSIONS

From Equation 3.13, the expected-return errors of the N assets are given by

$$e = E[R] - B\gamma. \quad (3.25)$$

A popular goodness-of-fit measure used in many empirical studies is the cross-sectional R^2 .

Following Kandel and Stambaugh (1995), this is defined as

$$R^2 = 1 - \frac{Q}{Q_0}, \quad (3.26)$$

where $Q = e'We$, $Q_0 = e_0'We_0$, $e_0 = [I_N - 1_N(1_N'W1_N)^{-1}1_N'W]E[R]$ represents the deviations of mean returns from their cross-sectional average, and W is a positive-definite weighting matrix.

Popular choices of W in the literature are $W = I_N$ (OLS), $W = Var[R]^{-1}$ (generalized least

squares, GLS), and $W = \Sigma_d^{-1}$ (weighted least squares, WLS), where Σ_d is a diagonal matrix containing the diagonal elements of Σ , the variance-covariance matrix of the residuals from the first-pass time series regression. In order for R^2 to be well defined requires assuming that $E[R]$ is not proportional to 1_N (the expected returns are not all equal) so that $Q_0 > 0$. Note that $0 < R^2 < 1$ and it is a decreasing function of the aggregate pricing-error measure Q . Thus, R^2 is a natural measure of goodness of fit.

As emphasized by Kan and Zhou (2004), R^2 is oriented toward expected returns whereas the HJ-distance evaluates a model's ability to explain prices. With the zero-beta rate as a free parameter, the most common approach in the asset pricing literature, they show that the two measures need not rank models the same way. Thus, both measures are of interest with the choice depending on the economic context and perhaps the manner in which a researcher envisions applying the models.

The estimated multiple regression betas of the N assets with respect to the K factors are defined as

$$\hat{\beta} = \hat{V}_{Rf} \hat{V}_f^{-1}, \quad (3.27)$$

where \hat{V}_{Rf} and \hat{V}_f are consistent estimators of $Cov[R, f]$ and $Var[f]$, respectively. Some studies allow β 's to change throughout the sample period. For example, in the original Fama and MacBeth (1973) study, the authors estimated the betas used in the CSR for month t from data before that month. A more customary practice in recent decades is to use full-period beta estimates for portfolios formed by ranking stocks on various characteristics.

Then, the estimated β 's are used as regressors in the second-pass CSR and the estimated risk premia, γ , are given by

$$\hat{\gamma} = (B_T' W_T B_T)^{-1} B_T' W_T \bar{R}, \quad (3.28)$$

where $B_T = [1_N, \hat{\beta}]$ and $\bar{R} = \frac{1}{T} \sum_{t=1}^T R_t$. Under the correctly specified model, the asymptotic standard errors of the risk premia estimates in Equation 3.28 are provided by Shanken (1992)

and Jagannathan and Wang (1998). Further, Shanken shows that, when the factors are portfolio returns, the most efficient estimates of the factor risk premia are the time-series means of the factors. He also shows how to incorporate the portfolio restriction in the cross-sectional relation when some of the factors are traded and others are not.

The vector of sample pricing errors is given by

$$\hat{e} = \bar{R} - B_T \hat{\gamma} \quad (3.29)$$

and the sample R^2 is

$$\hat{R}^2 = 1 - \frac{\hat{Q}}{\hat{Q}_0}, \quad (3.30)$$

where $\hat{Q} = \hat{e}' W_T \hat{e}$, $Q_0 = \hat{e}_0' W_T \hat{e}_0$, $\hat{e}_0 = [I_N - 1_N(1_N' W_T 1_N)^{-1} 1_N' W_T] \bar{R}$. To determine whether the model is correctly specified, one can test if the CSR R^2 is equal to one. Kan et al. (2010) show that if expected returns are exactly linear in the betas, then the limiting distribution of $T(\hat{R}^2 - 1)$ is that of a linear combination of $N - K - 1$ independent χ_1^2 random variables. Further, they characterize the asymptotic distribution of the sample R^2 when the true R^2 is 0 (i.e., the model has no explanatory power for expected returns) and when the true R^2 is between 0 and 1 (i.e., a misspecified model that provides some explanatory power for the expected returns on the test assets). Shanken (1985), Gibbons, Ross, and Shanken (1989), and Kan et al. (2010) provide alternative tests of the validity of the beta-pricing relation.

When the beta-pricing model is misspecified, the asymptotic standard errors proposed by Shanken (1992) and Jagannathan and Wang (1998) are incorrect. Shanken and Zhou (2007) and Kan et al. (2010) show how to compute misspecification-robust standard errors of the risk premia estimates. Finally, Kan et al. (2010) provide the necessary econometric techniques to compare the cross-sectional R^2 s of two or more beta-pricing models. As for the HJ-distance measure, the asymptotic distributions of their tests depend on whether the models are correctly specified or misspecified, nested or non-nested.

CONDITIONAL ASSET PRICING MODELS AND RETURN PREDICTABILITY

Recall that the fundamental pricing equation in Equation 3.2 is defined in terms of conditional expectations. Although the law of iterated expectations permits the estimation of the model in terms of unconditional moments, some relevant information may get lost in the process. This section explains how to incorporate conditioning information in a linear asset pricing model, describes the underlying assumptions, and provides an interpretation of the zero-beta rate and risk premia in the cross-sectional regression.

Let z_t be an L -vector of observed conditioning variables (instruments) that belongs to the information set at time t and define $F_{t+1} = [z'_t, f'_{t+1}, z'_t \otimes f'_{t+1}]'$ as a $\tilde{K} = (K + 1)(L + 1) - 1$ vector of scaled factors. Recently, many empirical studies (see, for example, Shanken, 1990; Lettau and Ludvigson, 2001; Lustig and Van Nieuwerburgh, 2005; Santos and Veronesi, 2006) have considered a cross-sectional regression of unconditional expected returns on their unconditional betas with respect to F_{t+1} :

$$E[R_{t+1}] = 1_N \gamma_0 + \beta \gamma_1, \quad (3.31)$$

where

$$\beta = Cov[R_{t+1}, F_{t+1}] Var[F_{t+1}]^{-1}. \quad (3.32)$$

There are two ways for obtaining the unconditional relationship between $E[R_{t+1}]$ and β in Equation 3.31. The first approach is a time-varying SDF coefficients approach, which assumes that the SDF is linear in a set of risk variables:

$$m_{t+1} = a_t + b'_t f_{t+1}. \quad (3.33)$$

The linearity of the SDF in f_{t+1} allows obtaining the following conditional asset pricing model:

$$E[R_{t+1}|z_t] = 1_N \gamma_{0,t} + \beta_t \gamma_{1,t}, \quad (3.34)$$

where $\beta_t = Cov[R_{t+1}, f_{t+1}|z_t] Var[f_{t+1}|z_t]^{-1}$ is the matrix of conditional betas and

$$\gamma_{0,t} = \frac{1}{E[m_{t+1}|z_t]} - 1, \quad (3.35)$$

$$\gamma_{1,t} = - \frac{Var[f_{t+1}|z_t] b_t}{E[m_{t+1}|z_t]}. \quad (3.36)$$

If a zero-beta asset exists with raw return $R_{0,t+1}$ (with $R_{0,t+1}$ conditionally uncorrelated with m_{t+1}), it follows that

$$\gamma_{0,t} = 1 + E[R_{0,t+1}|z_t]. \quad (3.37)$$

The SDF coefficients, a_t and b_t , are often assumed to be linear functions of the instruments z_t :

$$a_t = a_0 + a'_1 z_t \quad (3.38)$$

and

$$b_t = b_0 + B_1 z_t. \quad (3.39)$$

Then, m_{t+1} can be written as

$$m_{t+1} = a_0 + \tilde{b}' F_{t+1}, \quad (3.40)$$

where $\tilde{b} = [a'_1, b'_0, \text{vec}(B_1)']'$. As a result, assuming that the coefficients of the SDF are linear in z_t is equivalent to assuming that the SDF is linear in the scaled factors. For example, in a model with one risk factor and one conditioning variable, the SDF is given by

$$m_{t+1} = (a_0 + a_1 z_t) + (b_0 + b_1 z_t) f_{t+1} = a_0 + a_1 z_t + b_0 f_{t+1} + b_1 (z_t f_{t+1}), \quad (3.41)$$

i.e., going from a one-factor model with time-varying coefficients to a three-factor model with fixed coefficients is possible. Therefore, one can use the new (scaled) factors with the unconditional moment procedure developed above.

However, γ_0 in Equation 3.31 should not be interpreted as the unconditional expected return on the zero-beta asset. The reason is that, using Equation 3.37, Jensen's inequality and the fact that $E[m_{t+1}|z_t]$ is a positive random variable (positivity of m_{t+1} is required here), it follows that

$$E[R_{0,t+1}] \geq \gamma_0. \quad (3.42)$$

The equality holds if and only if $E[m_{t+1}|z_t]$ is constant over time. This result suggests that if a risk-free asset exists, then γ_0 tends to be less than the average risk-free rate. Similarly, the elements of γ_1 that correspond to the original K factors should not be interpreted as unconditional risk premia.

The second approach that can deliver Equation 3.31 is a time-varying regression coefficients approach (Shanken, 1990) with data generating process given by

$$R_{t+1} = \alpha_t + \beta_t f_{t+1} + \varepsilon_{t+1}, \quad (3.43)$$

where β_t is the matrix of conditional betas defined above and $\alpha_t = E[R_{t+1}|z_t] - \beta_t E[f_{t+1}|z_t]$.

When the conditional K -factor beta-pricing model holds,

$$E[R_{t+1}|z_t] = 1_N \gamma_{0,t} + \beta_t \gamma_{1,t}. \quad (3.44)$$

Assuming the conditional expectation of Equation 3.43 holds, the conditional K -factor beta-pricing model imposes the following restrictions on α_t :

$$\alpha_t = 1_N \gamma_{0,t} + \beta_t (\gamma_{1,t} - E[f_{t+1}|z_t]) = 1_N \gamma_{0,t} + \beta_t \varphi_t, \quad (3.45)$$

where $\varphi_t = \gamma_{1,t} - E[f_{t+1}|z_t]$. The betas, the zero-beta rate, and the risk premia in Equation 3.45 are time-varying. Since these restrictions are too general to test, some ancillary assumptions are needed. One possibility is to assume that $\gamma_{0,t}$ and φ_t are constant over time and that α_t and β_t are linear functions of z_t :

$$\alpha_t = a_0 + A_1 z_t, \quad (3.46)$$

$$\text{vec}(\beta_t) = b_0 + B_1 z_t. \quad (3.47)$$

The restriction in Equation 3.45 then becomes:

$$a_0 + A_1 z_t = 1_N \gamma_0 + (\varphi' \otimes I_N)(b_0 + B_1 z_t). \quad (3.48)$$

This restriction implies that

$$a_0 = 1_N \gamma_0 + (\varphi' \otimes I_N) b_0 \quad (3.49)$$

and

$$A_1 = (\varphi' \otimes I_N) B_1. \quad (3.50)$$

The return generating process can be written as

$$R_{t+1} = a_0 + A_1 z_t + (f'_{t+1} \otimes I_N)(b_0 + B_1 z_t) + \varepsilon_{t+1}, \quad (3.51)$$

which has $N(K+1)(L+1)$ parameters. This result is the same as running a regression of R_{t+1} on a constant term and the scaled factors F_{t+1} :

$$R_{t+1} = \tilde{\alpha} + \tilde{\beta}_1 z_t + \tilde{\beta}_2 f_{t+1} + \tilde{\beta}_3 (z_t \otimes f_{t+1}) + \varepsilon_{t+1}. \quad (3.52)$$

Comparing Equations 3.51 and 3.52 yields

$$\tilde{\alpha} = a_0, \quad \tilde{\beta}_1 = A_1, \quad \text{vec}(\tilde{\beta}_2) = b_0, \quad \text{vec}(\tilde{\beta}_3) = \text{vec}(B_1). \quad (3.53)$$

Taking the unconditional expectation of Equation 3.52,

$$E[R_{t+1}] = \tilde{\alpha} + \tilde{\beta}_1 E[z_t] + \tilde{\beta}_2 E[f_{t+1}] + \tilde{\beta}_3 E[z_t \otimes f_{t+1}]. \quad (3.54)$$

Using the asset pricing restrictions

$$\tilde{\alpha} = 1_N \gamma_0 + \tilde{\beta}_2 \varphi, \quad (3.55)$$

it follows that

$$E[R_{t+1}] = 1_N \gamma_0 + \tilde{\beta}_1 E[z_t] + \tilde{\beta}_2 (\varphi + E[f_{t+1}]) + \tilde{\beta}_3 E[z_t \otimes f_{t+1}]. \quad (3.56)$$

The risk premia associated with $\tilde{\beta}$ in the regression setup are not free, which differs from the time-varying SDF coefficients setup. In the regression framework, γ_0 indeed has a zero-beta interpretation and the risk premia associated with $\tilde{\beta}_2$ are indeed equal to the risk premia on the original factors. However, this relationship comes at the expense of assuming that γ_0 and φ are constant over time. Additionally, imposing the restriction $A_1 = (\varphi' \otimes I_N) B_1$, which is equivalent to $\tilde{\beta}_1 = \tilde{\beta}_3 (I_L \otimes \varphi)$, results in a simpler cross-sectional regression:

$$E[R_{t+1}] = 1_N \gamma_0 + \tilde{\beta}_2 (\varphi + E[f_{t+1}]) + \tilde{\beta}_3 E[z_t \otimes (\varphi + f_{t+1})]. \quad (3.57)$$

In summary, both approaches can lead to the unconditional relationship in Equation 3.31. However, the time-varying regression coefficients approach requires many assumptions (linearity assumptions on $N(K+1)$ regression coefficients together with constant γ_0 and φ). In contrast, the SDF approach requires far fewer assumptions (linearity assumption on $K+1$ SDF coefficients) and does not assume that the zero-beta rate and the risk premia are constant over time. In the regression approach, the γ_0 and γ_1 associated with the original factors retain the zero-beta rate and risk premia interpretation. Conversely, in the SDF approach, the γ_0 and γ_1 associated with the original factors cannot be interpreted as unconditional zero-beta rate and risk premia. Finally, Equation 3.57 shows that under the regression approach, the cross-

sectional regression should be run with a constant, $\tilde{\beta}_2$ and $\tilde{\beta}_3$ only. Although the risk premia associated with $\tilde{\beta}_1$ are in general not zero (unless the information variables are de-meanned), including $\tilde{\beta}_1$ does not provide additional explanatory power in the cross-sectional regression.

Conditional asset pricing models presume the existence of some return predictability. For the conditional restriction in Equation 3.2 to be empirically relevant, there should exist some instruments z_t for which the first and second moments of the SDF and of the returns vary over time. A typical predictive regression model of stock returns has the form:

$$R_{t+1} = \alpha + z_t' \beta + e_{t+1}, \quad (3.58)$$

where e_{t+1} is a martingale difference sequence. The vector of financial and macro predictors z_t includes valuation ratios (dividend-price ratio, dividend yields, earnings-price ratio, dividend-earnings ratio, and book-to-market ratio), interest and inflation rates (short-term rates, yield spreads, default premium, and inflation rate), consumption and wealth-income ratio, stock return volatility (realized or implied volatility), among others.

The main drawback of this approach is the reliance on a small number of conditioning variables, which is unlikely to span the information set of market participants (Ludvigson and Ng, 2007). Furthermore, the predictive ability of individual conditioning variables, if there is any, is only short-lived (Timmermann, 2008), unstable and subject to structural breaks over longer time periods (Lettau and Van Nieuwerburgh, 2008). To a large extent, these drawbacks can be remedied by estimating a few common factors from a large panel of economic time series that are believed to span the information set of investors (Ludvigson and Ng, 2007).

To introduce the main idea behind the estimation of common factors, suppose that the researcher has access to a large panel of data x_{it} ($i = 1, \dots, M; t = 1, \dots, T$), where M is the number of variables (financial and macro variables) and T is the number of time series observations. Assume that x_{it} admits an approximate factor structure of the form:

$$x_{it} = \omega_i' f_t + e_{it}, \quad (3.59)$$

where f_t is a K -vector of latent factors, ω_i is a K -vector of factor loadings and e_{it} are errors uncorrelated with the factors. Let $X = [x_1 \ x_2 \ \dots \ x_M]$ and $F = [f_1 \ f_2 \ \dots \ f_K]$ denote the stacked matrices for the data and the factors. Then, under some technical conditions, the latent factors can be estimated by the method of principal components analysis by minimizing the objective function

$$\frac{1}{MT} \sum_{i=1}^M \sum_{t=1}^T (x_{it} - \omega_i' f_t)^2, \quad (3.60)$$

subject to the identifying restriction $\frac{F'F}{T} = I_K$. The problem of estimating f_t is identical to maximizing $\text{tr}(F'(X'X)F)$ and the estimated factors \hat{f}_t are \sqrt{T} times the K eigenvectors corresponding to the K largest eigenvalues of the matrix $XX'/(MT)$.

The evaluation of predictability of stock returns is performed either in-sample or out-of-sample using statistical or economic criteria. The in-sample predictability is assessed in terms of the time-series R^2 of the model and of the statistical significance of the coefficient on a particular predictor. Typically, predictive regressions of stock returns are characterized by a statistically small but possibly economically relevant R^2 (Campbell and Thompson, 2008). As discussed later, the statistical significance of the slope parameter may be misleading if the predictor is highly persistent.

The out-of-sample prediction is performed by dividing the sample into two subsamples with the first subsample used for parameter estimation and the second subsample used for out-of-sample forecast evaluation. The statistical evaluation is based on the out-of-sample R^2 coefficient, mean squared or absolute errors that compare the actual and predicted values of the returns. Conversely, the profit-based evaluation involves computing returns from a trading strategy of stocks and bonds depending on whether the predicted excess returns from the model are positive (position in stocks) or negative (position in bonds). Then, the Sharpe ratio of the model-based trading strategy is compared to the Sharpe ratio of a buy-and-hold benchmark strategy over the out-of-sample evaluation period. Welch and Goyal (2008) provide a

comprehensive study of the out-of-sample performance of various financial and macroeconomic variables for predicting stock returns.

NONLINEAR ASSET PRICING MODELS

The generality of the SDF representation and GMM estimation based on the HJ-distance becomes obvious in the case, for example, of nonlinear consumption-based asset pricing models. As discussed above, the SDF for a representative agent model can be written as the product between a time-preference parameter β and the ratio of the marginal utilities of consumption at time $t + 1$ and t , respectively.

Consider the constant relative risk aversion (CRRA) or power utility function

$$u(c_t) = \frac{c_t^{1-\rho} - 1}{1-\rho}, \quad (3.61)$$

where $\rho > 0$ is the coefficient of relative risk aversion. For example, when $\rho \rightarrow 1$, $u(c_t) = \log(c_t)$. The Arrow-Pratt coefficient of relative risk aversion $\frac{c_t u''(c_t)}{u'(c_t)}$ is ρ , i.e., relative risk aversion is constant.

Substituting into the fundamental pricing equation delivers the following set of moments:

$$E_t \left[\beta \left(\frac{c_{t+1}}{c_t} \right)^{-\rho} (1 + R_{t+1}) - 1_N \right] = 0_N. \quad (3.62)$$

For a vector of instruments (conditioning variables) z_t that belongs to the information set at time t , the sample analog of the above population moment condition is

$$g_T(\theta) = \frac{1}{T} \sum_{t=1}^T \left[\beta \left(\frac{c_{t+1}}{c_t} \right)^{-\rho} (1 + R_{t+1}) - 1_N \right] \otimes z_t = 0_m, \quad (3.63)$$

where $m = \dim(R_{t+1}) \dim(z_t)$ and $\theta = (\beta, \rho)'$. The unknown parameters θ are then estimated by GMM.

Several drawbacks of the CRRA utility function are worth mentioning. First, the equity premium puzzle (Mehra and Prescott, 1985) implies unrealistically large values of risk aversion (ρ as high as 30 to 50) in order to fit U.S. data. For example, in a gamble that offers a 50

percent chance to double one's wealth and 50 percent chance to cut one's wealth by half, a value of risk aversion parameter of 30 implies that one would be willing to pay 49 percent of her wealth to hedge against the 50 percent chance of losing half of her wealth (Siegel and Thaler, 1997). Second, in the CRRA framework, ρ is inversely related to the elasticity of intertemporal substitution (EIS). This is inappropriate because EIS is related to the willingness of an investor to transfer consumption between time periods whereas CRRA is about transferring consumption between states of the world. The non-expected and time non-separable (habit persistence) utility functions described below separate risk aversion and intertemporal substitution.

The non-expected (Epstein-Zin-Weil) utility is given by

$$u(c_t) = [(1 - \beta)c_t^{1-\eta} + \beta(E_t[u_{t+1}^{1-\rho}])^{(1-\eta)/(1-\rho)}]^{1/(1-\eta)} \quad (3.64)$$

which gives rise to the following pricing equation (conditional moment restriction):

$$E_t \left[\beta^\lambda \left(\frac{c_{t+1}}{c_t} \right)^{-\eta\lambda} (1 + R_{m,t+1})^{\lambda-1} (1 + R_{t+1}) - 1_N \right] = 0_N, \quad (3.65)$$

where $\lambda = \frac{1-\rho}{1-\eta}$ and R_m denotes the market return. Note that for $\lambda = 1$ (or, equivalently, $\eta = \rho$), this equation reduces to the one corresponding to time-separable (CRRA) utility. The sample analog of the moment condition above is given by

$$g_T(\theta) = \frac{1}{T} \sum_{t=1}^T \left[\beta^\lambda \left(\frac{c_{t+1}}{c_t} \right)^{-\eta\lambda} (1 + R_{m,t+1})^{\lambda-1} (1 + R_{t+1}) - 1_N \right] \otimes z_t = 0_m, \quad (3.66)$$

where the parameter vector is $\theta = (\beta, \eta, \lambda)'$.

Another popular extension of the CRRA framework is the utility function with habit persistence and durability:

$$u(c_t, c_{t-1}) = \frac{s_t^{1-\rho}}{1-\rho}, \quad (3.67)$$

where $s_t = c_t + \tau c_{t-1}$. The conditional moment restrictions are given by

$$E_t [\beta (s_{t+1}^{-\rho} + \beta \tau s_{t+2}^{-\rho}) (1 + R_{t+1}) - (s_t^{-\rho} + \beta \tau s_{t+1}^{-\rho})] = 0_N, \quad (3.68)$$

and their sample analog takes the form

$$g_T(\theta) = \frac{1}{T} \sum_{t=1}^T \left[\beta \left(\frac{s_{t+1}^{-\rho} + \beta \tau s_{t+2}^{-\rho}}{s_t^{-\rho}(1+\beta\tau)} \right) (1 + R_{t+1}) - \left(\frac{s_t^{-\rho} + \beta \tau s_{t+1}^{-\rho}}{s_t^{-\rho}(1+\beta\tau)} \right) 1_N \right] \otimes z_t = 0_m, \quad (3.69)$$

where $\theta = (\beta, \rho, \tau)'$ and the original moment conditions are divided by the $s_t^{-\rho}(1 + \beta\tau)$ habit to induce stationarity and rule out trivial solutions (note that if $\rho = 0$ and $\beta\tau = -1$, the moment conditions are trivially satisfied). As in the case of non-expected utility, the time-separable (CRRA) utility is a special case of Equation 3.69 for $\tau = 0$. While estimating the set of moment conditions for the habit persistence model and testing the implied model restrictions are possible, Equation 3.69 does not have a clear pricing error interpretation. One possibility is to make lognormality assumptions as in Balduzzi and Kallal (1997), cast the restrictions in pricing error form and use the HJ-distance metric for model evaluation and comparison. To conclude, all linear and nonlinear asset pricing models that have been proposed in the literature can be written in terms of the fundamental asset pricing equation in Equation 3.2 and can be estimated using GMM-type techniques.

PITFALLS IN THE CURRENT PRACTICE AND SUGGESTIONS FOR IMPROVING EMPIRICAL WORK

One empirical finding that consistently emerges from the statistical tests and comparisons of competing asset pricing models is that the data are too noisy for a meaningful and conclusive differentiation among alternative SDF specifications. Given the large noise component in returns on risky assets, explaining the cross-sectional variability of asset returns by using slowly changing financial and macroeconomic variables appears to be a daunting task. Even if the asset pricing theory provides guidance for the model specification, the properties of the data and some limitations of the standard statistical methodology can create further challenges in applied work. This section discusses several pitfalls that accompany the estimation of risk premia and evaluation of competing asset pricing models using actual data. Particular attention is paid to the possibility of model misspecification, presence of useless factors, highly persistent

conditioning variables, large number of test assets, potential lack of invariance to data scaling, and interpretation of the risk premia.

Misspecified Models

A widely-held belief is that asset pricing models are likely to be misspecified and should be viewed only as approximations of the true data generating process. Nevertheless, empirically evaluating the degree of misspecification and the relative pricing performance of candidate models using actual data is useful.

Two main problems with the econometric analyses are present when performed in the existing asset pricing studies. First, even when a model is strongly rejected by the data (using one of the model specification tests previously described, for example), researchers still construct standard errors of parameter estimates using the theory developed for correctly specified models. This process could give rise to highly misleading inference especially when the degree of misspecification is large. Kan and Robotti (2009) and Gospodinov et al. (2011a) focus on the HJ-distance metric and derive misspecification-robust standard errors of the SDF parameter estimates for linear and nonlinear models. In contrast, Kan et al. (2010) focus on the beta representation of an asset pricing model and propose misspecification-robust standard errors of the second-pass risk premia estimates. For example, for linear SDF specifications, the misspecification adjustment term, associated with the misspecification uncertainty surrounding the model, can be decomposed into three components: (1) a pure misspecification component that captures the degree of misspecification, (2) a spanning component that measures the degree to which the factors are mimicked by returns, and (3) a component that measures the usefulness of the factors in explaining the variation in returns. The adjustment term is zero if the model is correctly specified (component (1) is zero) and/or the factors are fully mimicked by returns (component (2) is zero). If the factors are poorly mimicked by the returns, the adjustment

term could be very large. This issue will be revisited in the discussion of the useless factors case in the next section.

Second, many researchers are still ranking competing models by simply eyeballing the differences in sample HJ-distances or sample R^2 's without any use of a formal statistical criterion that accounts for the sampling and model misspecification uncertainty. Kan and Robotti (2009), Kan et al. (2010), and Gospodinov et al. (2011a) develop a complete statistical procedure for comparing alternative asset pricing models. These model selection tests take into account the restrictions imposed by the structure of the competing models (nested, non-nested or overlapping) as well as the estimation and model misspecification uncertainty. Gospodinov et al. (2011a) also propose chi-squared versions of these tests that are easy to implement and enjoy excellent finite-sample properties.

One recommendation for empirical work that emerges from these remarks is that the statistical inference in asset pricing models should be conducted allowing for the possibility of potential misspecification. This will ensure robust and valid inference in the presence of model misspecification as well as when the models are correctly specified.

Useless Factors

Consistent estimation and valid inference in asset pricing models crucially depends on the identification condition that the covariance matrix of asset returns and risk factors is of full rank. Kan and Zhang (1999a, 1999b) study the consequences of the violation of this identification condition. In particular, they show that when the model is misspecified and one of the included factors is useless (i.e., independent of asset returns), the asymptotic properties of parameter and specification tests in GMM and two-pass cross-sectional regressions are severely affected.

The first serious implication of the presence of a useless factor is that the asymptotic distribution of the Wald test (squared t -test) of statistical significance of the useless factor's parameter (HJ-distance case) is chi-squared distributed with $N - K - 1$ degrees of freedom

instead of one degree of freedom as in the standard case when all factors are useful. The immediate consequence of this result is that the Wald test that uses critical values from a chi-squared distribution with one degree of freedom will reject the null hypothesis too frequently when the null hypothesis is true. The false rejections are shown to become more severe as the number of test assets N becomes large and as the length of the sample increases. As a result, researchers may erroneously conclude that the useless factor is priced when, in reality, it is pure noise uncorrelated with the stock market.

Another important implication is that the true risk premium associated with the useless factor is not identifiable and the estimate of this risk premium diverges at rate \sqrt{T} . The standard errors of the risk-premium estimates associated with the useful factors included in the model are also affected by the presence of a useless factor and the standard inference is distorted. Similar results also arise for optimal GMM estimation (Kan and Zhang, 1999a) and two-pass cross-sectional regressions (Kan and Zhang, 1999b).

The useless factor problem is particularly serious because the traditional model specification tests previously described cannot reliably detect misspecification in the presence of a useless factor. This manifests itself in the failure of the specification tests to reject the null hypothesis of correct specification when the model is indeed misspecified and contains a useless factor.

More generally, similar types of problems are symptomatic of a violation of the crucial identification condition that the covariance matrix of asset returns and risk factors must be of full rank. Therefore, a rank restriction test (see, for example, Gospodinov, Kan, and Robotti, 2010b) should serve as a useful pre-test for possible identification problems in the model (see also Burnside, 2010). However, this test cannot identify which factor contributes to the identification failure. Kleibergen (2009) proposes test statistics that exhibit robustness to the degree of correlation between returns and factors in a two-pass cross-sectional regression framework. In the SDF framework, Gospodinov, Kan, and Robotti (2011b) develop a simple (asymptotically

χ_1^2 -distributed) misspecification-robust test that signals the direction of the identification failure. Only after the useless factor is detected and removed from the analysis, the validity of the (misspecification-robust) inference and the consistency of the parameter estimates can be restored.

Estimating Models with Excess Returns

When excess returns (R^e) are used to estimate and test asset pricing models, the moment conditions (pricing equation) are given by

$$E(mR^e) = 0_N. \quad (3.70)$$

Let $m = \theta_0 - (\theta_1 f_1 + \dots + \theta_K f_K)$. In this case, the mean of the SDF cannot be identified or, equivalently, the parameters θ_0 and $(\theta_1, \dots, \theta_K)$ cannot be identified separately. This requires a particular choice of normalization. One popular normalization is to set $\theta_0 = 1$ in which case $m = 1 - (\theta_1 f_1 + \dots + \theta_K f_K)$. An alternative (preferred) normalization is to set $\theta_0 = 1 + \theta_1 E(f_1) + \dots + \theta_K E(f_K)$ in which case $m = 1 - \theta_1 [f_1 - E(f_1)] - \dots - \theta_K [f_K - E(f_K)]$ with $E(m) = 1$. These two normalizations can give rise to very different results (see Kan and Robotti, 2008; Burnside, 2010).

Kan and Robotti (2008) argue that when the model is misspecified, the first (raw) and the second (de-meanned) normalizations of the SDF produce different GMM estimates that minimize the quadratic form of the pricing errors. Hence, the pricing errors and the p -values of the specification tests are not identical under these two normalizations. Moreover, the second (de-meanned) specification imposes the constraint $E(m) = 1$ and, as a result, the pricing errors and the HJ-distances are invariant to affine transformations of the factors. This is important because in the first normalization, the outcome of the model specification test can be easily manipulated by simple scaling of factors and changing the mean of the SDF. This problem is not only a characteristic of linear SDFs but also arises in nonlinear models. The analysis in Burnside (2010) further confirms these findings and links the properties of the different normalizations to

possible model misspecification and identification problems discussed in the previous two subsections.

In a two-pass CSR framework, Kan et al. (2010) explore an excess returns specification with the zero-beta rate constrained to equal the risk-free rate. Imposing this restriction seems sensible since, when the beta-pricing models are estimated with the zero-beta rate as a free parameter, the estimated zero-beta rate is often too high and the estimated market premium is often negative, contrary to what economic theory suggests. The zero-beta restriction in the CSR context can be implemented by working with test portfolio returns in excess of the T-bill rate, while excluding the constant from the expected return relations. As is typical for regression analysis without a constant, the corresponding R^2 measure involves (weighted) sums of squared values of the dependent variable (mean excess returns) in the denominator, not squared deviations from the cross-sectional average.

With the zero-beta rate constrained in this manner, it follows from the results of Kan and Robotti (2008) that equality of GLS R^2 's for two models is equivalent to equality of their HJ-distances, provided that the SDF is written as a linear function of the de-meaned factors as mentioned above. No such relation exists for the OLS R^2 .

Conditional Models with Highly Persistent Predictors

The usefulness of the conditional asset pricing models crucially depends on the existence of some predictive power of the conditioning variables for future stock returns. While a large literature reports statistically significant coefficients for various financial and macro variables in in-sample linear predictive regressions of stock returns, several papers raise the concern that some of these regressions may be spurious. For example, Ferson, Sarkissian, and Simin (2003) call into question the predictive power of some widely used predictors such as the term spread, book-to-market ratio, and dividend yield. Spurious results arise when the predictors are strongly persistent (near unit root processes) and their innovations are highly correlated with the

predictive regression errors. In this case, the estimated slope coefficients in the predictive regression are biased and have a non-standard (non-normal) asymptotic distribution (Elliott and Stock, 1994; Cavanagh, Elliott, and Stock, 1995; Stambaugh, 1999). As a result, t -tests for statistical significance of individual predictors based on standard normal critical values could reject the null hypothesis of no predictability too frequently and falsely signal that these predictors have predictive power for future stock returns. Campbell and Yogo (2006) and Torous, Valkanov, and Yan (2004) develop valid testing procedures when the predictors are highly persistent and revisit the evidence on predictability of stock returns.

Spuriously significant results and non-standard sampling distributions also tend to arise in long-horizon predictive regressions where the regressors and/or the returns are accumulated over r time periods so that two or more consecutive observations are overlapping. The time overlap increases the persistence of the variables and renders the sampling distribution theory of the slope coefficients, t -tests and R^2 coefficients, non-standard. Campbell (2001) and Valkanov (2003) point out several problems that emerge in long-horizon regressions with highly persistent regressors. First, the R^2 coefficients and t -statistics tend to increase with the horizon, even under the null of no predictability, and the R^2 is an unreliable measure of goodness of fit in this situation. Furthermore, the t -statistics do not converge asymptotically to well-defined distributions and need to be rescaled to ensure valid inference. Finally, the estimates of the slope coefficients are biased and, in some cases, not consistently estimable. All these statistical problems provide a warning to applied researchers and indicate that the selection of conditioning variables for predicting stock returns should be performed with extreme caution.

Model Evaluation with a Large Number of Assets

A common practice is to evaluate the empirical relevance of different asset pricing models using a relatively large cross-section of returns on 25, 50 or 100 portfolios at monthly or quarterly frequencies over a period of 30 years (Fama and French, 1992; Jagannathan and Wang, 1996).

Since the size of the cross-section, N , determines the dimensionality of the vector of moment conditions in the GMM estimation, the small number of time series observations per moment condition renders the asymptotic approximations for some specification and model comparison tests inaccurate.

For instance, Ahn and Gadarowski (2004) report substantial size distortions of the specification test based on the HJ-distance for combinations of N and T that are typically encountered in practice. In particular, in some of their simulation designs, the HJ-distance test rejects the null hypothesis, when the null hypothesis is true, 99 percent (51 percent) of the time for $N = 100$ and $T = 160$ ($T = 330$) at the 1 percent nominal level. This indicates that the researcher will erroneously conclude with high probability that the asset pricing model under investigation is misspecified. These simulation results suggest that the weighted chi-squared asymptotic approximation (for N fixed and T approaching infinity) is inappropriate when the number of test assets is large.

Several testing procedures for correct model specification with improved finite-sample properties are available in the literature. Kan and Zhou (2004) derive the exact distribution of the sample HJ-distance which can be obtained by simulation. On the other hand, Gospodinov et al. (2011a) continue to use the weighted chi-squared asymptotic approximation but compute the weights for this asymptotic distribution not from the variance matrix of the pricing errors (moment conditions) computed under the null hypothesis but from its analog computed under the alternative of misspecification. While these variance matrices are asymptotically equivalent under the null of correct specification, the variance matrix computed under the alternative tends to be larger in finite samples, thus rendering the too frequent rejection problem less severe. Finally, Gospodinov et al. (2011a) propose an alternative model specification test that measures the distance of the Lagrange multipliers, associated with the pricing constraints imposed by the model, from zero. This new test is easy to implement (critical values are based on a chi-squared

distribution with $N - K - 1$ degrees of freedom) and is characterized by excellent size and power properties.

Beta or Covariance Risk?

Historically, researchers have almost exclusively focused on the price of beta risk to infer whether a proposed factor is priced. However, a potential issue exists with using multiple regression betas when $K > 1$: in general, the beta of an asset with respect to a particular factor depends on what other factors are included in the first-pass time-series OLS regression. As a consequence, the interpretation of the risk premia given in Equation 3.28 in the context of model selection can be problematic. For example, suppose that a model has two factors f_1 and f_2 . Interest often lies in determining whether f_2 is needed in the model. Some researchers have tried to answer this question by performing a test of $H_0: \gamma_2 = 0$, where γ_2 is the risk premium associated with factor 2. When the null hypothesis is rejected by the data, they typically conclude that factor 2 is important, and when the null hypothesis is not rejected, they conclude that factor 2 is unimportant.

Kan et al. (2010) provide numerical examples illustrating that the test of $H_0: \gamma_2 = 0$ does not answer the question of whether factor 2 helps to explain the cross-sectional differences in expected returns on the test assets. They also provide two solutions to this problem. The first remedy is to use simple regression betas instead of multiple regression betas in the second-pass CSR. The second solution consists in running the second-pass CSR with covariances instead of betas. Kan and Robotti (2011) derive the asymptotic theory for the case of simple-regression betas, while Kan et al. (2010) provide inference techniques for second-pass regressions that are run with covariances instead of betas. Therefore, researchers should focus on the price of covariance risk and not on the price of beta risk. Finding a statistically significant price of covariance risk is indeed evidence that the underlying factor is incrementally useful in explaining the cross-section of asset returns.

SUMMARY AND CONCLUSIONS

This chapter provides an up-to-date review of the two most popular approaches for estimating, testing and comparing potentially misspecified asset pricing models: the stochastic discount factor and the beta methods. The analysis points out various pitfalls with some popular usages of these methodologies that could lead to erroneous conclusions. Special emphasis is given to the role played by model misspecification in tests of unconditional and conditional asset pricing models, to the important issue of selecting information variables that truly predict future returns and to different ways of incorporating the predictions of asset pricing theory into competing empirical specifications.

Although the recommendations in this chapter are specifically designed to sharpen asset pricing tests and provide a bigger challenge to the existing models, much remains to be done. On the one hand, given the limited number of time-series observations for stocks and bonds, the asymptotic methods summarized in this chapter should be complemented with more reliable finite-sample procedures. Conversely, whether researchers should use individual assets or aggregated portfolios in tests of asset pricing theories is not entirely clear. Although the finance profession seems to favor the idea of working with portfolios instead of individual assets, justifying the almost exclusive reliance on the 25 size and book-to-market Fama-French portfolio returns is difficult. How many portfolios should be considered and how should they be formed are certainly open questions that future research will hopefully address.

DISCUSSION QUESTIONS

1. Discuss the advantages and the drawbacks of the Hansen-Jagannathan distance and cross-sectional R^2 for evaluating and comparing possibly misspecified asset pricing models.
2. Some studies suggest that the predictive power of different financial and macro variables for forecasting future stock returns should be evaluated only out-of-sample, i.e., using

information only up to the time when the forecast is made. List several reasons that could justify the preference for out-of-sample over in-sample evaluation of predictive power.

3. The SDF approach discussed in this chapter can be used for evaluating the performance of mutual and hedge funds. Describe briefly how the SDF approach can be implemented in practice for this task if mutual/hedge fund data are available.
4. Despite the recent developments in asset pricing theory and practice, many statistical problems can still potentially compromise some empirical findings reported in the literature. Discuss some of the pitfalls in the empirical analysis of asset pricing models.

REFERENCES

- Abel, Andrew B. 1990. "Asset Prices under Habit Formation and Catching Up with the Joneses." *American Economic Review* 80:2, 38–42.
- Ahn, Seung C., and Christopher Gadarowski. 2004. "Small Sample Properties of the Model Specification Test Based on the Hansen-Jagannathan Distance." *Journal of Empirical Finance* 11:1, 109–132.
- Balduzzi, Pierluigi, and Hedi Kallal. 1997. "Risk Premia and Variance Bounds." *Journal of Finance* 52:5, 1913–1949.
- Bansal, Ravi, and Amir Yaron. 2004. "Risks for the Long-Run: A Potential Resolution of Asset Pricing Puzzles." *Journal of Finance* 59:4, 1481–1509.
- Berkman, Henk, Ben Jacobsen, and John B. Lee. 2011. "Time-Varying Rare Disaster Risk and Stock Returns." *Journal of Financial Economics* 101:2, 313–332.
- Black, Fischer, Michael C. Jensen, and Myron Scholes. 1972. "The Capital Asset Pricing Model: Some Empirical Tests." In Michael C. Jensen, ed., *Studies in the Theory of Capital Markets*, 79–121. New York: Praeger.
- Breeden, Douglas T. 1979. "An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities." *Journal of Financial Economics* 7:3, 265–296.

- Burnside, Craig. 2010. "Identification and Inference in Linear Stochastic Discount Factor Models." NBER Working Paper No. 16634.
- Campbell, John Y. 2001. "Why Long Horizons? A Study of Power against Persistent Alternatives." *Journal of Empirical Finance* 8:5, 459–491.
- Campbell, John Y., and John H. Cochrane. 1999. "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior." *Journal of Political Economy* 107:2, 205–251.
- Campbell, John Y., and John H. Cochrane. 2000. "Explaining the Poor Performance of Consumption-Based Asset Pricing Models." *Journal of Finance* 55:6, 2863–2878.
- Campbell, John Y., and Samuel B. Thompson. 2008. "Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average?" *Review of Financial Studies* 21:4, 1509–1531.
- Campbell, John Y., and Motohiro Yogo. 2006. "Efficient Tests of Stock Return Predictability." *Journal of Financial Economics* 81:1, 27–60.
- Cavanagh, Christopher L., Graham Elliott, and James H. Stock. 1995. "Inference in Models with Nearly Integrated Regressors." *Econometric Theory* 11:5, 1131–1147.
- Chen, Xiaohong, and Sydney C. Ludvigson. 2009. "Land of Addicts? An Empirical Investigation of Habit-Based Asset Pricing Models." *Journal of Applied Econometrics* 24:7, 1057–1093.
- Cochrane, John H. 2005. *Asset Pricing*. Princeton, NJ: Princeton University Press.
- Constantinides, George M. 1990. "Habit Formation: A Resolution of the Equity Premium Puzzle." *Journal of Political Economy* 98:3, 519–543.
- Dittmar, Robert F. 2002. "Nonlinear Pricing Kernels, Kurtosis Preference, and Evidence from the Cross-Section of Equity Returns." *Journal of Finance* 57:1, 369–403.
- Elliott, Graham, and James H. Stock. 1994. "Inference in Time Series Regression when the Order of Integration of a Regressor is Unknown." *Econometric Theory* 10:3–4, 672–700.

- Epstein, Larry G., and Stanley E. Zin. 1989. "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework." *Econometrica* 57:4, 937–968.
- Epstein, Larry G., and Stanley E. Zin. 1991. "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Investigation." *Journal of Political Economy* 99:2, 555–576.
- Fama, Eugene F. 1991. "Efficient Capital Markets: II." *Journal of Finance* 46:5, 1575–1617.
- Fama, Eugene F., and Kenneth R. French. 1992. "The Cross-Section of Expected Stock Returns." *Journal of Finance* 47:2, 427–465.
- Fama, Eugene F., and Kenneth R. French. 1993. "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics* 33:1, 3–56.
- Fama, Eugene F., and James D. MacBeth. 1973. "Risk, Return, and Equilibrium: Empirical Tests." *Journal of Political Economy* 81:3, 607–636.
- Farnsworth, Heber, Wayne E. Ferson, David Jackson, and Steven Todd. 2002. "Performance Evaluation with Stochastic Discount Factors." *Journal of Business* 75:3, 473–503.
- Ferson, Wayne E., and George M. Constantinides. 1991. "Habit Persistence and Durability in Aggregate Consumption." *Journal of Financial Economics* 29:2, 199–240.
- Ferson, Wayne E., Sergei Sarkissian, and Timothy T. Simin. 2003. "Spurious Regressions in Financial Economics?" *Journal of Finance* 58:4, 1393–1413.
- Gibbons, Michael R., Stephen A. Ross, and Jay Shanken. 1989. "A Test of the Efficiency of a Given Portfolio." *Econometrica* 57:5, 1121–1152.
- Gospodinov, Nikolay, Raymond Kan, and Cesare Robotti. 2010a. "On the Hansen-Jagannathan Distance with a No-Arbitrage Constraint." Federal Reserve Bank of Atlanta Working Paper 2010-4.

- Gospodinov, Nikolay, Raymond Kan, and Cesare Robotti. 2010b. "Further Results on the Limiting Distribution of GMM Sample Moment Conditions." Federal Reserve Bank of Atlanta Working Paper 2010-11.
- Gospodinov, Nikolay, Raymond Kan, and Cesare Robotti. 2011a. "Chi-Squared Tests for Evaluation and Comparison of Asset Pricing Models." Federal Reserve Bank of Atlanta Working Paper 2011-8.
- Gospodinov, Nikolay, Raymond Kan, and Cesare Robotti. 2011b. "Solving the Useless Factor Problem in Linear Asset Pricing Models." Unpublished manuscript.
- Hansen, Lars Peter. 1982. "Large Sample Properties of Generalized Method of Moments Estimators." *Econometrica* 50:4, 1029–1054.
- Hansen, Lars Peter, and Ravi Jagannathan. 1997. "Assessing Specification Errors in Stochastic Discount Factor Models." *Journal of Finance* 52:2, 557–590.
- Hodrick, Robert J., and Xiaoyan Zhang. 2001. "Evaluating the Specification Errors of Asset Pricing Models." *Journal of Financial Economics* 62:2, 327–376.
- Jagannathan, Ravi, Keiichi Kubota, and Hitoshi Takehara. 1998. "Relationship between Labor-Income Risk and Average Return: Empirical Evidence from the Japanese Stock Market." *Journal of Business* 71:3, 319–348.
- Jagannathan, Ravi, Georgios Skoulakis, and Zhenyu Wang. 2010. "The Analysis of the Cross-Section of Security Returns." In Yacine Ait-Sahalia and Lars Hansen, eds., *Handbook of Financial Econometrics*, Volume 2, Chapter 14, 73–134. Amsterdam: Elsevier.
- Jagannathan, Ravi, and Zhenyu Wang. 1996. "The Conditional CAPM and the Cross-Section of Expected Returns." *Journal of Finance* 51:1, 3–53.
- Jagannathan, Ravi, and Zhenyu Wang. 1998. "An Asymptotic Theory for Estimating Beta-Pricing Models Using Cross-Sectional Regression." *Journal of Finance* 53:4, 1285–1309.
- Kan, Raymond, and Cesare Robotti. 2008. "Specification Tests of Asset Pricing Models Using Excess Returns." *Journal of Empirical Finance* 15:5, 816–838.

- Kan, Raymond, and Cesare Robotti. 2009. "Model Comparison Using the Hansen-Jagannathan Distance." *Review of Financial Studies* 22:9, 3449–3490.
- Kan, Raymond, and Cesare Robotti. 2011. "On the Estimation of Asset Pricing Models Using Univariate Betas." *Economics Letters* 110:2, 117–121.
- Kan, Raymond, Cesare Robotti, and Jay Shanken. 2010. "Pricing Model Performance and the Two-Pass Cross-Sectional Regression Methodology." Unpublished manuscript.
- Kan, Raymond, and Chu Zhang. 1999a. "GMM Tests of Stochastic Discount Factor Models with Useless Factors." *Journal of Financial Economics* 54:1, 103–127.
- Kan, Raymond, and Chu Zhang. 1999b. "Two-Pass Tests of Asset Pricing Models with Useless Factors." *Journal of Finance* 54:1, 204–235.
- Kan, Raymond, and Guofu Zhou. 2004. "Hansen-Jagannathan Distance: Geometry and Exact Distribution." Unpublished manuscript.
- Kandel, Shmuel, and Robert F. Stambaugh. 1995. "Portfolio Inefficiency and the Cross-Section of Expected Returns." *Journal of Finance* 50:1, 157–184.
- Kleibergen, Frank. 2009. "Tests of Risk Premia in Linear Factor Models." *Journal of Econometrics* 149:2, 149–173.
- Lettau, Martin, and Sydney C. Ludvigson. 2001. "Resurrecting the (C)CAPM: A Cross-Sectional Test When Risk Premia are Time-Varying." *Journal of Political Economy* 109:6, 1238–1287.
- Lettau, Martin, and Stijn Van Nieuwerburgh. 2008. "Reconciling the Return Predictability Evidence." *Review of Financial Studies* 21:4, 1607–1652.
- Li, Haitao, Yewu Xu, and Xiaoyan Zhang. 2010. "Evaluating Asset Pricing Models Using the Second Hansen-Jagannathan Distance." *Journal of Financial Economics* 97:2, 279–301.
- Lintner, John. 1965. "The Valuation of Risky Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets." *Review of Economics and Statistics* 47:1, 13–37.
- Long, John Jr. 1974. "Stock Prices, Inflation, and the Term Structure of Interest Rates." *Journal of Financial Economics* 1:2, 131–170.

- Ludvigson, Sydney C., and Serena Ng. 2007. "The Empirical Risk-Return Relation: A Factor Analysis Approach." *Journal of Financial Economics* 83:1, 171–222.
- Lustig, Hanno, and Stijn Van Nieuwerburgh. 2005. Housing Collateral, Consumption Insurance, and Risk Premia." *Journal of Finance* 60:3, 1167–1221.
- Mehra, Rajnish, and Edward C. Prescott. 1985. "The Equity Premium: A Puzzle." *Journal of Monetary Economics* 15:2, 145–161.
- Merton, Robert C. 1973. "An Intertemporal Capital Asset Pricing Model." *Econometrica* 41:5, 867–887.
- Parker, Jonathan A., and Christian Julliard. 2005. "Consumption Risk and the Cross-Section of Expected Returns." *Journal of Political Economy* 113:1, 185–222.
- Rubinstein, Mark. 1976. "The Valuation of Uncertain Income Streams and the Pricing of Options." *Bell Journal of Economics* 7:2, 407–425.
- Santos, Tano, and Pietro Veronesi. 2006. "Labor Income and Predictable Stock Returns." *Review of Financial Studies* 19:1, 1–44.
- Shanken, Jay. 1985. "Multivariate Tests of the Zero-Beta CAPM." *Journal of Financial Economics* 14:3, 327–348.
- Shanken, Jay. 1990. "Intertemporal Asset Pricing: An Empirical Investigation." *Journal of Econometrics* 45:1-2, 99–120.
- Shanken, Jay. 1992. "On the Estimation of Beta-Pricing Models." *Review of Financial Studies* 5:1, 1–33.
- Shanken, Jay, and Guofu Zhou. 2007. "Estimating and Testing Beta Pricing Models: Alternative Methods and their Performance in Simulations." *Journal of Financial Economics* 84:1, 40–86.
- Sharpe, William. 1964. "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk." *Journal of Finance* 19:3, 425–442.

- Siegel, Jeremy J., and Richard H. Thaler. 1997. "Anomalies: The Equity Premium Puzzle." *Journal of Economic Perspectives* 11:1, 191–200.
- Stambaugh, Robert F. 1999. "Predictive Regressions." *Journal of Financial Economics* 54:3, 375–421.
- Timmermann, Allan. 2008. "Elusive Return Predictability." *International Journal of Forecasting* 24:1, 1–18.
- Torous, Walter, Rossen Valkanov, and Shu Yan. 2004. "On Predicting Stock Returns with Nearly Integrated Explanatory Variables." *Journal of Business* 77:4, 937–966.
- Valkanov, Rossen. 2003. "Long-Horizon Regressions: Theoretical Results and Applications." *Journal of Financial Economics* 68:2, 201–232.
- Weil, Philippe. 1989. "The Equity Premium Puzzle and the Risk-Free Rate Puzzle." *Journal of Monetary Economics* 24:3, 401–421.
- Welch, Ivo, and Amit Goyal. 2008. "A Comprehensive Look at the Empirical Performance of Equity Premium Prediction." *Review of Financial Studies* 21:4, 1455–1508.

ABOUT THE AUTHORS

Nikolay Gospodinov is an Associate Professor of Economics at Concordia University and a Research Fellow of CIREQ. His research focuses on issues in econometrics, asset pricing, and international finance. He has published one book and more than 20 refereed articles in journals such as the *Review of Economics and Statistics*, *Journal of Business and Economic Statistics*, *Econometric Theory*, *Journal of Econometrics*, *Econometric Reviews*, *Journal of Financial Econometrics*, and *Journal of Futures Markets*. He is currently an Associate Editor of *Econometric Reviews*. He holds a Ph.D. in Economics from Boston College, and a B.A. and an M.A. in Economics from the University of National and World Economy.

Cesare Robotti is a financial economist and Associate Policy Adviser with the financial group of the research department of the Federal Reserve Bank of Atlanta and an Affiliate Professor at

EDHEC Business School. Dr. Robotti concentrates his research on empirical asset pricing and portfolio management. Before joining the Atlanta Fed in 2001, he was a teaching fellow at Boston College and an adjunct faculty member at Brandeis University. Dr. Robotti has also worked in the dealing room of Novara International Bank in Luxembourg and as a financial analyst at Cariplo Bank in Milan, Italy. He has published works in several leading academic journals and has presented papers and served as a discussant at various professional conferences. He holds a Ph.D. in Economics from Boston College, and a B.A. and an M.A. in Economics from L. Bocconi University.