Risk Premia and Variance Bounds
Author(s): Pierluigi Balduzzi and Hedi Kallal
Source: *The Journal of Finance*, Vol. 52, No. 5 (Dec., 1997), pp. 1913-1949
Published by: Blackwell Publishing for the American Finance Association
Stable URL: http://www.jstor.org/stable/2329469

THE JOURNAL OF FINANCE • VOL. LII, NO. 5 • DECEMBER 1997

Risk Premia and Variance Bounds

PIERLUIGI BALDUZZI and HÉDI KALLAL*

ABSTRACT

If a pricing kernel assigns a premium to a risk variable that differs from the one assigned by the minimum-variance admissible kernel, then the pricing kernel must exhibit more variability than the minimum-variance kernel. Based on this intuition, we derive a variance bound that is more stringent than that of Hansen and Jagannathan (1991). When we apply our bound to the kernel of a representative consumer with power utility, we find that the consumption risk premium increases the severity of the "equity-premium puzzle" of Mehra and Prescott (1985).

THERE IS CONSIDERABLE EVIDENCE against the Capital Asset Pricing Model (CAPM), suggesting that the CAPM fails to capture priced risks adequately. This has led several authors to investigate the empirical performance of the intertemporal CAPM (for example, see Chen, Roll, and Ross (1986), Ferson and Harvey (1991), and Fama and French (1993)) or multi-factor Arbitrage Pricing Theory (APT) models with prespecified variables (for example, see Burmeister and McElroy (1988)). This article offers a novel way of testing whether the premia associated with various sources of risk are consistent with the implications of security returns for specific asset pricing models. Specifically, we obtain a lower bound on the variance of an admissible pricing kernel for a given risk premium it assigns to an economic risk variable. We also obtain an upper and a lower bound on the risk premium assigned by a pricing kernel, for a given variance of the pricing kernel.

These bounds are closely related and extend the Hansen and Jagannathan (1991) variance bounds. They are obtained exploiting the following intuition: If a pricing kernel assigns a premium to a risk variable that differs from the one assigned by the minimum-variance admissible kernel, then the pricing kernel must exhibit more variability than the minimum-variance kernel. The variance of the minimum-variance kernel is the Hansen and Jagannathan variance bound. Moreover, the variance shown by a pricing kernel constrains the

* Balduzzi is from the Carroll School of Management, Boston College. Kallal is from Salomon Brothers. We thank Giorgio De Santis, John Heaton, Silverio Foresi, Ravi Jagannathan, Peter Knez, Anthony Lynch, Margaret McMillan, Ken Singleton, René Stulz, two anonymous referees, and seminar participants at the 1996 Winter Meetings of the Econometric Society, the 1996 Colloquia on Economic Research (IGIER, Milan), the 1996 Meetings of the Western Finance Association, and the 1997 Utah Winter Finance Conference for useful comments. We acknowledge financial support from New York University Summer Research Grants. This research was initiated while the authors were affiliated with New York University. Salomon is not responsible for any statement or conclusions herein, and no opinions, theories, or techniques presented herein in any way represent the position of Salomon Brothers, Inc.

way it interacts with a risk variable, and also the risk premium it assigns to that variable. We identify upper and lower bounds for the risk premium assigned to the risk variable. The interval created by these upper and lower bounds increases in size with the variance of a pricing kernel and also with the variance of the risk variable.

Our approach can be better understood in the context of some recent multifactor models. For example, Fama and French (1993) explain the premia on stocks by their exposure to the returns on three "factor" portfolios: a "market-factor" portfolio, a "size-factor" portfolio, and a "book-to-market-value-factor" portfolio. Following Hansen and Jagannathan (1991) we construct the minimum variance pricing kernel consistent with these returns. We show that if the representative investor has power utility and risk aversion such that the Hansen and Jagannathan bound is satisfied, then she assigns a risk premium to consumption growth that is significantly higher than the risk premium assigned by the three factor model. This difference in risk premia means that the consumption based kernel needs to be even more volatile than the kernel associated with the three factor model. Hence, using the Hansen and Jagannathan bound a researcher might erroneously conclude that the properties of a pricing kernel are consistent with asset returns when they are not.

Our analysis adds several features to the examination of risk premia. First, the validity of our bounds does not depend on the appropriate specification of either an asset-pricing model or of a stochastic process for asset returns. Second, our risk variables do not need to be spanned by traded asset returns. Third, while we allow for time variation of the risk premia assigned by a pricing kernel, time variation does not have to be explicitly modeled to calculate our bounds. Fourth, we can introduce conditioning information along the lines of Hansen and Richard (1987) to expand effectively the set of assets under scrutiny and sharpen our bounds.

As an application, our bounds on the risk premia and variance of a pricing kernel can be used to diagnose an explicit asset-pricing model. A candidate pricing kernel can in fact have a variance that is consistent with the Hansen and Jagannathan (1991) variance bound. Still, we show that the risk premia that the candidate kernel assigns to economic risk variables generate sharper variance bounds. These sharper bounds could lead to the rejection of the model. Moreover, if a candidate pricing kernel *does not* satisfy Hansen and Jagannathan's variance bound, we might be tempted to add some variability to it, in order to meet the restriction. Our variance bounds show that this does not necessarily lead to a model consistent with the data on economic sources of risk. Indeed, as the variability of the candidate pricing kernel increases, the difference likely increases between the premium that it assigns to a risk variable and the premium assigned by the minimum-variance kernel. This makes our variance bounds even more stringent.

Our bounds also apply in situations in which we have a priori information on variance that allows us to place bounds on the pricing kernel's risk premia. The reason for obtaining bounds on risk premia, as opposed to point estimates, is that we typically observe an incomplete collection of assets, and not all risk premia can be directly estimated from such data.

For the empirical analysis, we consider the standard pricing kernels of a representative consumer. The consumer has constant relative risk aversion, with preferences that are either separable or exhibit habit persistence over consumption at different times. We calculate our variance bounds for five economic variables. These variables proxy for corporate default, term structure slope, inflation, industrial production, and consumption-growth risk. We then test whether the candidate consumption-based pricing kernel meets the Hansen and Jagannathan (1991) variance bound; whether our variance bounds are more stringent than the Hansen and Jagannathan bound; and whether the candidate pricing kernel meets our variance bounds.

We find that our variance bounds are significantly more stringent than the Hansen and Jagannathan (1991) bound for both inflation and consumption risk. We also find that our variance bounds are more stringent from an economic standpoint, since they imply increases of 16.38 percent, for inflation risk and 71.28 percent for consumption risk, in the coefficient of relative risk aversion consistent with variance implications. We also find that the study of consumption risk leads to a marginal rejection of the consumption-based model, even for levels of relative risk aversion as high as 110 in the time-separable case, and as high as 12 in the habit-persistence case.

This last result exacerbates the "equity-premium puzzle" of Mehra and Prescott (1985). It also contrasts with Cochrane and Hansen (1992), who find that a relative risk aversion of one (log utility) does not lead to a strong rejection of the consumption-based model, based on quarterly returns on stocks and bonds. And it contrasts with Cecchetti, Lam, and Mark (1994), who also find that a relative risk aversion coefficient near one is marginally consistent with monthly data on stock and bond returns.

Several recent articles have applied the methodology of Hansen and Jagannathan (1991) to the testing of asset pricing models (Cecchetti, Lam, and Mark (1994) and Cochrane and Hansen (1992), for example). Other articles have tested whether different sets of securities generate significantly different minimum variance kernels (for example, see Chen and Knez (1994), De Santis (1995), Downs and Snow (1994), Knez (1994), and Snow (1991)). Our article is the first to show how risk variables in addition to asset returns can be used to make the Hansen and Jagannathan bounds more stringent, and to use these new bounds to test an explicit asset pricing model. Our article is also the first to show how the variance of a pricing kernel determines an upper and a lower bound on the values that a risk premium might take.

The article is organized as follows: Section I discusses the economic significance of a pricing kernel's cross moments. Section II derives the variance and risk premium bounds and shows how the bounds can be made sharper by including conditioning information and imposing the positivity restriction on a pricing kernel. Section III discusses how asset-pricing models can be tested, based on our variance bounds. Section IV considers a consumption-based pricing kernel, various economic risk variables, and a set of asset returns. Section V presents the results of the tests. Section VI concludes.

I. Cross Moments and Risk Premia

This section derives a general pricing result: The expected excess return on any asset depends on the covariance of the asset return with a pricing kernel. We then introduce a set of mean-zero, orthogonal risk variables.

Consider the case where we can construct portfolios that exactly replicate the risk variables. In this case the cross moments of the mimicking portfolio returns with the pricing kernel determine the expected cash flows on zero-net-investment positions long in the mimicking portfolio and short in the riskless asset.

A. A General Pricing Result

We denote by r the $N \times 1$ vector of (gross) security returns r under consideration. By the law of one price, we have

$$\mathrm{E}_t(m_{t+1} r_{t+1}) = \mathbf{1} \tag{1}$$

for some admissible pricing kernel m, where "$\mathbf{1}$" denotes an $N \times 1$ vector of ones.[1] Also, note that r can be measured either in units of the consumption basket or in units of the currency. Accordingly, m can be interpreted as either a real or a nominal pricing kernel.

We assume that the vector r includes the possibly time-varying rate of return on a risk-free asset, r_{ft}. This leads to

$$\mathrm{E}_t(m_{t+1} r_{ft}) = 1. \tag{2}$$

It is convenient to perform the analysis that follows in terms of the pricing kernel scaled by the risk-free rate, the normalized pricing kernel, $m_{t+1} r_{ft} \equiv q_{t+1}$, where $\mathrm{E}_t(q_{t+1}) = 1$.

Using equations (1) and (2), we obtain the familiar orthogonality condition

$$\mathrm{E}_t[q_{t+1}(r_{t+1} - \mathbf{1} r_{ft})] = 0. \tag{3}$$

Rearranging, we obtain

$$\mathrm{E}_t(q_{t+1})[E_t(r_{t+1}) - \mathbf{1} r_{ft}] = \mathrm{E}_t(r_{t+1}) - \mathbf{1} r_{ft} = -\mathrm{Cov}_t(q_{t+1}, r_{t+1}). \tag{4}$$

This means that the conditional expected return on any asset in excess of the risk-free rate equals the negative of the conditional covariance between the asset return and the normalized pricing kernel.

[1] Harrison and Kreps (1979) provide the relevant discussion. The set of pricing kernels can be interpreted as the set of intertemporal marginal rates of substitution compatible with the distribution of returns.

B. Risk Variables and Mimicking Portfolios

We denote by y the $K \times 1$ vector of K risk variables under consideration, $y_{k,t+1}$, $k = 1, \ldots, K$. Without loss of generality, we assume these variables to be orthogonal to each other and to have a constant zero mean and a unit variance, i.e.,

$$\mathrm{E}_t(y_{k,t+1}) = \mathrm{E}_t(y_{k,t+1}y_{j,t+1}) = 0, \quad \text{and} \quad \mathrm{E}_t(y_{k,t+1}^2) = 1 \tag{5}$$

for any $k, j = 1, \ldots, K$. The reason for the orthogonality assumption is that it lets us attribute variability to the different risk variables in an unambiguous way. The reason for assuming a constant zero mean is that the only variability "priced" by security markets is that which is not anticipated. The relevant risk variables are innovations in economic variables. The reason for scaling the risk variables with a constant unit variance is that the resulting risk premia are in the units of the traditional Sharpe ratio: Expected excess return per unit of risk.

Assume for now that there exist portfolios that exactly mimic the behavior of the risk variables. Let

$$r_{y_k,t+1} \equiv \mathrm{E}_t(r_{y_k,t+1}) + y_{k,t+1} \tag{6}$$

denote the returns on the exact mimicking portfolios. According to equation (4), we have

$$\mathrm{E}_t(r_{y_k,t+1}) - r_{ft} = -\mathrm{Cov}_t(q_{t+1}, r_{y_k,t+1}) = -\mathrm{E}_t(q_{t+1}y_{k,t+1}) \equiv \lambda_{kt} \tag{7}$$

which is the conditional premium on the corresponding risk variable, y_k. This means that the conditional cross moment between the normalized pricing kernel q and the risk variable y_k equals, with the opposite sign, the conditional risk premium on y_k. Assuming stationarity and applying the law of iterated expectations, we obtain

$$\lambda_k \equiv \mathrm{E}(\lambda_{kt}) \equiv -\mathrm{E}(qy_k) \tag{8}$$

which is the unconditional or mean risk premium on y_k. This means that the unconditional cross moment between the normalized pricing kernel q and the risk variable y_k equals, with the opposite sign, the mean risk premium on y_k.

II. Risk Premia and Variance Bounds

We now make inferences on risk premia without explicit assumptions on the existence of exact mimicking portfolios, or on the form of the pricing kernel and the return-generation process. The approach that we take here is similar in spirit to Hansen and Jagannathan (1991), henceforth H&J.

First, we derive the minimum-variance bounds for an admissible normalized pricing kernel. Next, we calculate the mean risk premia assigned by the minimum-variance kernel, and we show how they correspond to the mean cash flows generated by the (approximate) mimicking portfolios financed at the riskless rate. We show that if the risk premia assigned by an admissible (normalized) pricing kernel differ from those assigned by the minimum-variance kernel, then the variance of the kernel must strictly exceed that of the minimum-variance kernel. Conversely, we show how our analysis can be used to derive bounds on risk premia, given a priori knowledge of the variance of the pricing kernel.

A. The Hansen and Jagannathan Variance Bound

Using the definition of the normalized pricing kernel q, the pricing equation (1) can be rearranged into

$$\mathrm{E}_t(q_{t+1} r_{t+1}) = \mathbf{1} r_{ft}. \tag{9}$$

This means that all expected asset returns, after a risk adjustment, should equal the risk-free rate of return. We denote by $r_a \equiv [r^\top, 1]^\top$ the augmented vector of asset returns and by $\iota \equiv [\mathbf{1}^\top r_f, 1]^\top$ the augmented riskless-rate vector. Assuming stationarity and applying the law of iterated expectations to equation (9) and to the condition $\mathrm{E}_t(q_{t+1}) = 1$ we obtain

$$\mathrm{E}(q r_a) = \mathrm{E}(\iota). \tag{10}$$

Following H&J, we can construct an admissible (normalized) pricing kernel—i.e., a random variable that satisfies (10)—that is linear in r_a: $q^* \equiv r_a^\top \alpha$, where α is an $(N + 1) \times 1$ coefficient vector. We must have

$$\mathrm{E}[(r_a^\top \alpha) r_a] = \mathrm{E}(\iota). \tag{11}$$

Assuming $\mathrm{E}(r_a r_a^\top)$ to be invertible, we obtain $\alpha = [\mathrm{E}(r_a r_a^\top)]^{-1}\mathrm{E}(\iota)$. Note that $\mathrm{E}(q^*) = \mathrm{E}(q) = 1$. Also, we have $\mathrm{E}[r_a(q - q^*)] = 0$, since both q and q^* satisfy (10). Hence, $(q - q^*)$ is orthogonal to r_a and q^* is the fitted value of a theoretical regression of q on r_a. We have $\mathrm{Var}(q) = \mathrm{Var}(q^*) + \mathrm{Var}(q - q^*)$, and hence[2]

$$\mathrm{Var}(q) \geq \mathrm{Var}(q^*), \qquad \mathrm{E}(q) = \mathrm{E}(q^*) = 1. \tag{12}$$

B. Mimicking Portfolios and Risk Premia

Now consider the fitted value of a theoretical regression of the risk variable y_k on r_a, $y_k^* \equiv r_a^\top \alpha_{y_k}$, where α_{y_k} is an $(N + 1) \times 1$ coefficient vector. The first N

[2] Also related to the analysis of Hansen and Jagannathan (1991) is MacKinlay (1995) who relates deviations from the CAPM to the minimum variability of an admissible pricing kernel, and tests whether these deviations can be attributed to missing risk factors or to nonrisk-based sources.

elements of α_{y_k} represent the amounts that must be invested in the N assets to best replicate the variability of the risk variable y_k. Note that the mimicking portfolio does not need to be exact and that in general, y_k differs from y_k^*. We then have

$$\mathrm{E}[(y_k - (r_a^\top \alpha_{y_k}))r_a] = 0. \tag{13}$$

This implies $\alpha_{y_k} = [\mathrm{E}(r_a r_a^\top)]^{-1}\mathrm{E}(r_a y_k)$. Note that since $\mathrm{E}[r_a(y_k - y_k^*)] = 0$, we also have $\mathrm{E}(y_k - y_k^*) = 0$ and $\mathrm{E}(y_k) - \mathrm{E}(y_k^*) = 0$, due to the fact that r_a contains a unit payoff. Hence, $\mathrm{E}(y_k^*) = \mathrm{E}(y_k) = 0$.

Consider now the cross moments between the minimum-variance kernel q^* and the risk variables y_k. Let $\lambda_k^* \equiv -\mathrm{E}(qy_k^*) = -\mathrm{E}(q^*y_k^*)$. Since q^* satisfies the unconditional moment restriction (10) by construction, we have

$$\lambda_k^* = -\mathrm{E}(q^*y_k^*) = -\mathrm{E}[q^*(\alpha_{y_k}^\top r_a)] = -\alpha_{y_k}^\top \mathrm{E}(\iota). \tag{14}$$

The result above can be restated as

$$\lambda_k^* = -\alpha_{y_k 1}\mathrm{E}(r_f) - \alpha_{y_k 2}\mathrm{E}(r_f) - \cdots - \alpha_{y_k N}\mathrm{E}(r_f) - \alpha_{y_k, N+1}. \tag{15}$$

On the other hand, we have

$$y_k^* = \alpha_{y_k 1} r_1 + \alpha_{y_k 2} r_2 + \cdots + \alpha_{y_k N} r_N + \alpha_{y_k, N+1}. \tag{16}$$

Taking expected values, we obtain

$$\alpha_{y_k 1}\mathrm{E}(r_1) + \alpha_{y_k 2}\mathrm{E}(r_2) + \cdots + \alpha_{y_k N}\mathrm{E}(r_N) + \alpha_{y_k, N+1} = 0 \tag{17}$$

since $\mathrm{E}(y_k^*) = 0$. Using (15) and (17), we obtain

$$\lambda_k^* = \alpha_{y_k 1}\mathrm{E}(r_1 - r_f) + \alpha_{y_k 2}\mathrm{E}(r_2 - r_f) + \cdots + \alpha_{y_k N}\mathrm{E}(r_N - r_f) \tag{18}$$

which is the mean cash flow from being long in the portfolio best mimicking the risk variable y_k and short in the riskless asset. In other words, λ_k^* is the mean risk premium on the best approximate mimicking portfolio for the risk variable y_k. This implies that λ_k^* does *not* depend on the choice of the (normalized) pricing kernel, but only on the asset returns under scrutiny.

C. Risk Premia and Variance Bounds

Here, we do not assume the existence of exact mimicking portfolios. Hence risk premia may not be estimated directly from security-market data. Instead, we assign values to the risk premia λ_k (defined in equation (8)), and we derive the corresponding implications for the variance of a normalized pricing kernel. This approach complements that of H&J, who assign values to the (assumed) unobservable mean price of the riskless asset.

The mean risk premium λ_k can be written as:

$$\lambda_k \equiv -\mathrm{E}(qy_k) \tag{19}$$

$$= -\mathrm{E}[(q^* + (q - q^*))(y_k^* + (y_k - y_k^*))] \tag{20}$$

$$= -\mathrm{E}(q^*y_k^*) - \mathrm{E}[q^*(y_k - y_k^*)] - \mathrm{E}[y_k^*(q - q^*)] - \mathrm{E}[(q - q^*)(y_k - y_k^*)]. \tag{21}$$

Since both $(y_k - y_k^*)$ and $(q - q^*)$ are orthogonal to r_a, we have $\mathrm{E}[q^*(y_k - y_k^*)] = \mathrm{E}[(q - q^*)y_k^*] = 0$, and hence $\lambda_k = \lambda_k^* - \mathrm{E}[(q - q^*)(y_k - y_k^*)]$. Rearranging, we obtain

$$\mathrm{E}[(q - q^*)(y_k - y_k^*)] = \lambda_k^* - \lambda_k. \tag{22}$$

Moreover, by the Cauchy-Schwarz inequality, we have

$$|\mathrm{E}[(q - q^*)(y_k - y_k^*)]| \leq \sqrt{\mathrm{E}[(q - q^*)^2]\mathrm{E}[(y_k - y_k^*)^2]}. \tag{23}$$

Since $\mathrm{E}(q) = \mathrm{E}(q^*) = 1$, we also have $\mathrm{E}(q - q^*)^2 = \mathrm{E}(q^2) - \mathrm{E}[(q^*)^2] = \mathrm{Var}(q) - \mathrm{Var}(q^*)$. (Recall that q is equal to q^* plus an error term uncorrelated with q^*). Taking the square of both sides of equation (23), using equation (22), and rearranging, we obtain

$$\mathrm{Var}(q) \geq \mathrm{Var}(q^*) + \frac{[\mathrm{E}((q - q^*)(y_k - y_k^*))]^2}{\mathrm{E}[(y_k - y_k^*)^2]} = \mathrm{Var}(q^*) + \frac{(\lambda_k^* - \lambda_k)^2}{\mathrm{E}[(y_k - y_k^*)^2]}. \tag{24}$$

Note that the right-hand side of equation (24) as a function of λ_k describes a parabola, with a minimum of $\mathrm{Var}(q^*)$ at $\lambda_k = \lambda_k^*$.

The inequality in equation (24) shows that if the components of the pricing kernel q, and the components of the risk variable y_k that are not explained by the asset returns r exhibit covariability, the variance of q must exceed the variance of q^*. The lower bound on $\mathrm{Var}(q) - \mathrm{Var}(q^*)$ increases with $(\lambda_k^* - \lambda_k)^2$ and decreases with $\mathrm{E}[(y_k - y_k^*)^2]$ (where $0 \leq \mathrm{E}[(y_k - y_k^*)^2] \leq \mathrm{E}(y_k^2) = 1$), the variance of y_k that is not explained by asset returns.

Note that $(\lambda_k^* - \lambda_k)^2$ is not a measure of mispricing. Indeed, since the (normalized) pricing kernel, q, is assumed to correctly price the security returns under scrutiny, r, and to have unit mean, then by construction we have $\lambda_k^* = -\mathrm{E}(qy_k^*)$. On the other hand, the (normalized) pricing kernel q can covary with risk variables that are orthogonal to r_a, and hence can assign risk premia that differ from those assigned by q^*. This means that in general, we have $\lambda_k \neq \lambda_k^*$. In fact, the quantity $(\lambda_k - \lambda_k^*)/\sqrt{\mathrm{E}[(y_k - y_k^*)^2]}$ is the Sharpe ratio of a position that is long in the risk variable y_k and short in its best approximate mimicking portfolio. In other words, it is the amount by which an investor with preferences described by the (normalized) pricing kernel q needs to be compensated (on average) to bear one unit of the risk that is associated with the risk variable y_k and that is not captured by the asset returns r.

We see the intuition behind the variance bounds in equation (24) most clearly in the case in which y_k is orthogonal to the (augmented) space of asset returns. In this case, we have $\mathrm{Var}(q) \geq \mathrm{Var}(q^*) + \lambda_k^2$. In particular, if we think of y_k as a "new" asset return, that is "added" to the existing set of asset returns r, then the term λ_k^2, which is (the square of) the risk premium on y_k, is the upward shift in the H&J variance bound that is due to the introduction of the new asset return.[3] Also especially interesting is when the (normalized) pricing kernel q is spanned by the (augmented) set of security returns r_a. In this case, our variance bounds reduce to $\mathrm{Var}(q) = \mathrm{Var}(q^*)$. For example, if the pricing kernel q were a linear function of some asset returns and $q = q^*$, we would have $\lambda_k = \lambda_k^*$.

Also note that the inequality in equation (24) can be extended to the case in which several risk variables are considered simultaneously. Let q_y^* denote the minimum-variance random variable such that $-\mathrm{E}[q_y^*(y - y^*)] = \lambda - \lambda^*$, where y and y^* are the vectors of y_k and y_k^*, respectively, and λ and λ^* are the vectors of risk premia λ_k and λ_k^*, respectively. We know that $q_y^* = (y - y^*)^\top[\mathrm{E}((y - y^*)(y - y^*)^\top)]^{-1}(\lambda - \lambda^*)$. Therefore, we also have

$$\mathrm{Var}(q_y^*) = (\lambda - \lambda^*)^\top[\mathrm{E}((y - y^*)(y - y^*)^\top)]^{-1}(\lambda - \lambda^*) = \sum_{k=1}^{K} \frac{(\lambda_k^* - \lambda_k)^2}{\mathrm{E}[(y_k - y_k^*)^2]}. \tag{25}$$

Consider now the admissible (normalized) pricing kernel q, which assigns the risk premia λ_k. Since $-\mathrm{E}(qy^*) = -\mathrm{E}(q^*y^*) = \lambda^*$, we have $-\mathrm{E}[q(y - y^*)] = \lambda_k - \lambda_k^*$. Hence, $\mathrm{Var}(q^*) + \mathrm{Var}(q_y^*)$ is a lower bound on the variance of q, which translates into the inequality

$$\mathrm{Var}(q) \geq \mathrm{Var}(q^*) + \sum_{k=1}^{K} \frac{(\lambda_k^* - \lambda_k)^2}{\mathrm{E}[(y_k - y_k^*)^2]}. \tag{26}$$

The variance bounds in equations (24) and (26) could be used as diagnostics of an explicit asset-pricing model. Indeed, when testing a pricing kernel, we do not have to limit ourselves to the first and second moments of the kernel and the behavior of asset returns, as in H&J. For instance, Snow (1991) examines the higher-order norms of a pricing kernel. Also, Cochrane (1992) focuses on the implications of the behavior of price-dividend ratios for the mean and variance of a stochastic discount factor.

We can further highlight the difference between our variance bounds and H&J's approach by using the following example: Consider a candidate (normalized) pricing kernel x that is implied by an asset-pricing model that we want to test. Assume that the candidate kernel does not satisfy the H&J

[3] We thank John Heaton for having suggested this intuition.

variance bound $\mathrm{Var}(x) < \mathrm{Var}(q^*)$. We could argue that by simply adding to x a random variable orthogonal to x and r_a with enough variance, we could obtain a new candidate kernel that meets the restriction. Our variance bounds make it difficult to fix the pricing kernel. In fact, as the variability of the candidate kernel increases, the difference between the risk premium that it assigns to a risk variable, $\lambda_{xk} \equiv -\mathrm{E}(xy_k)$, and the risk premium assigned by the minimum-variance kernel, λ_k^*, is likely to increase. This makes our variance bounds even more stringent. It also makes it more difficult for the new kernel to meet our restriction.

D. Risk-Premium Bounds

The bounds expressed in equation (24) place restrictions on the variance of an admissible (normalized) pricing kernel. These restrictions depend on the risk premia that a kernel assigns to the risk variables. Alternatively, our analysis can be used to place restrictions on the unobservable risk premia. These restrictions depend on the variance of a pricing kernel. In general, we would expect the size of the admissible interval for a risk premium to increase as the variance of a pricing kernel increases. Indeed, in this case, the kernel may exhibit a wide range of covariances with the risk variable.

More precisely, from (22) we have $\mathrm{E}[(q - q^*)(y_k - y_k^*)] = \lambda_k - \lambda_k^*$. Moreover, we have $\mathrm{E}(q - q^*)^2 = \mathrm{Var}(q) - \mathrm{Var}(q^*)$. Substituting in the inequality of equation (23), we obtain

$$\lambda_k^* - \sqrt{[\mathrm{Var}(q) - \mathrm{Var}(q^*)]\mathrm{E}[(y_k - y_k^*)^2]} \leq \lambda_k$$
$$\leq \lambda_k^* + \sqrt{[\mathrm{Var}(q) - \mathrm{Var}(q^*)]\mathrm{E}[(y_k - y_k^*)^2]}. \quad (27)$$

The size of the interval for λ_k increases with the variance of the normalized pricing kernel in excess of $\mathrm{Var}(q^*)$. It also increases with the variance of the risk variable y_k not explained by the asset returns r.

The economic message of the bounds expressed in equation (27) is as follows: Consider a situation in which we do not directly observe a pricing kernel. If we have a priori information as to its variability, we can still place bounds on the premia it assigns to any source of risk.

Note that the closer the risk variables are to being "traded," the tighter the bounds. If the assets allow the construction of a mimicking portfolio, then the risk variable y_k is spanned by asset returns, and the bounds (27) reduce to $\lambda_k = \lambda_k^*$. Instead of an admissible interval, we obtain a point estimate of the mean risk premium. At the other extreme, if $y_k^* = 0$, i.e., if the risk variable y_k is orthogonal to the augmented asset-return space, the bounds (27) simplify to[4] $|\lambda_k| \leq \sqrt{\mathrm{Var}(q) - \mathrm{Var}(q^*)}$.

[4] Remember that $\mathrm{E}(y_k^2) = 1$ by assumption, while, in this case, $\lambda_k^* \equiv -\mathrm{E}(q^*y_k^*) = 0$.

The discussion above highlights the reason for having bounds on the risk premia λ_k, rather than exact values. Typically we observe an incomplete collection of assets, which makes it impossible to hedge all sources of risk.[5]

E. Conditioning Information

While the analysis above focuses on the unconditional moment restriction (10), there is a simple way to incorporate conditional information along the lines indicated by Hansen and Richard (1987).[6]

Consider a vector of instruments z_t, whose realization belongs to the information set at time t.
Let

$$r^z_{a,t+1} \equiv [r^\top_{a,t+1} \otimes z^\top_t, z^\top_t]^\top \tag{28}$$

and

$$\iota^z_t \equiv [(\mathbf{1} \otimes z_t)^\top r^\top_{ft}, z^\top_t]^\top. \tag{29}$$

Our variance bounds can be reformulated by replacing r_a and ι with their scaled counterparts r^z_a and ι^z. (For details see the Appendix.)

Note that the scaled returns, $r^z_{a,t+1}$, are interpreted as cash flows generated by managed portfolios. Investors who observe z_t can invest in an asset according to the value of z_t. The theoretical regressions of the pricing kernel q and the risk variables y_k on the scaled returns r^z_a, rather than on r_a, are likely to capture more of their variability, and to make the bounds on the variance of the pricing kernel and the risk premia more stringent.

F. Positivity of the Pricing Kernel

While the minimum-variance pricing kernel q^* satisfies the law of one price, in general it does not satisfy the other no-arbitrage condition $q^* > 0$. Nonetheless, as in H&J, we can extend our analysis to take this restriction into account.

Let $\tilde{\alpha}$ denote an $(N + 1) \times 1$ coefficient vector, and define $\tilde{q} \equiv (r^\top_a \tilde{\alpha})^+ \equiv \max\{r^\top_a \tilde{\alpha}, 0\}$. Assume

$$\mathrm{E}(\tilde{q} r_a) = \mathrm{E}(\iota). \tag{30}$$

The random variable $\tilde{q}$ has the smallest variance among all non-negative random variables satisfying restriction (30). (H&J prove the existence of such

[5] For example, Cochrane and Saá-Requejo (1996) have independently developed a similar approach to calculate bounds on the price of a stock option when markets are incomplete.

[6] Hansen and Richard reinterpret the instrumental variables of Hansen and Singleton (1982) as conditional portfolio quantities.

a random variable.) In fact, we have $\mathrm{E}(qr_a) = \mathrm{E}(\tilde{q}r_a) = \mathrm{E}(\iota)$. Indeed, since q is nonnegative, we have

$$\mathrm{E}(\tilde{q}q) \geq \tilde{\alpha}^{\top}\mathrm{E}(r_a q) = \tilde{\alpha}^{\top}\mathrm{E}(r_a \tilde{q}) = \mathrm{E}(\tilde{q}^2). \tag{31}$$

Moreover, by the Cauchy-Schwarz inequality, we have $\mathrm{E}(q^2) \geq [\mathrm{E}(\tilde{q}q)]^2/\mathrm{E}(\tilde{q}^2)$, which, combined with the inequality (31), leads to

$$\mathrm{Var}(q) \geq \mathrm{Var}(\tilde{q}), \qquad \mathrm{E}(q) = \mathrm{E}(\tilde{q}) = 1. \tag{32}$$

Our variance and risk premium bounds can be reformulated as follows (the Appendix provides details): We have

$$\mathrm{Var}(q) \geq \mathrm{Var}(\tilde{q}) + \frac{(\tilde{\lambda}_k - \lambda_k)^2}{\mathrm{E}[(y_k - y_k^*)^2]} \tag{33}$$

where $\tilde{\lambda}_k \equiv -\mathrm{E}(\tilde{q}y_k)$. This is the analog of inequality (24). In the case of several risk variables we have

$$\mathrm{Var}(q) \geq \mathrm{Var}(\tilde{q}) + \sum_{k=1}^{K} \frac{(\tilde{\lambda}_k - \lambda_k)^2}{\mathrm{E}[(y_k - y_k^*)^2]} \tag{34}$$

which is the analog of equation (26).

We also have

$$\tilde{\lambda}_k - \sqrt{[\mathrm{Var}(q) - \mathrm{Var}(\tilde{q})]\mathrm{E}[(y_k - y_k^*)^2]} \leq \lambda_k$$
$$\leq \tilde{\lambda}_k + \sqrt{[\mathrm{Var}(q) - \mathrm{Var}(\tilde{q})]\mathrm{E}[(y_k - y_k^*)^2]} \tag{35}$$

which is the analog of the bounds (27) on the premia assigned to the risk variables, when the positivity restriction is not imposed. Note that the size of the admissible interval for the risk premium λ_k implied by inequality (35) is smaller than that implied by its counterpart (27), for any given value of $\mathrm{Var}(q)$.

If we combine the analysis of this section with that of Section II.E, which shows how conditional information could be incorporated, we obtain bounds on the variance of the pricing kernel and the risk premia. The variance is bounded under the positivity condition on the admissible pricing kernel, and with the asset returns scaled by conditioning variables.

III. Testing the Variance Bounds

In this section, we illustrate tests of a pricing kernel based on the inequalities expressed in equations (12) and (32) (the H&J variance bound), and (24) and (33) (our variance bounds). Specifically, we test three hypotheses. First, we use the H&J bound to test the candidate pricing kernel implied by an explicit asset pricing model. Second, we determine whether our variance bounds are statistically more stringent than the H&J bound. Third, we test the candidate pricing kernel using our more stringent variance bounds.

Let x denote a candidate (normalized) pricing kernel. We determine whether the candidate kernel x satisfies the H&J bound by testing the null hypothesis

$$\Delta^{\mathrm{HJ}} \equiv \operatorname{Var}(x) - \operatorname{Var}(q^*) \geq 0. \tag{36}$$

A negative and significant estimate of Δ^{HJ} indicates that the candidate kernel x fails to satisfy the H&J bound. This test is analogous to that performed by Cecchetti, Lam, and Mark (1994). One can also see also Cochrane and Hansen (1992), who compare the second moment of the candidate pricing kernel to that of the minimum-variance admissible kernel. Burnside (1994) finds that the small sample properties of these test statistics of the H&J regions are quite good.

Also, let $\lambda_{xk} \equiv -\mathrm{E}(xy_k)$ denote the premium on the risk variable y_k assigned by the candidate kernel. We determine whether our variance bounds are more stringent than the H&J bound by testing the null hypothesis

$$\delta_k \equiv \frac{\lambda_k^* - \lambda_{xk}}{\sqrt{\mathrm{E}[(y_k - y_k^*)^2]}} = 0. \tag{37}$$

Note that the alternative hypothesis $\delta_k \neq 0$ implies $\delta_k^2 = (\lambda_k^* - \lambda_{xk})^2/\mathrm{E}[(y_k - y_k^*)^2] > 0$. Hence, a rejection of the null hypothesis in this equation implies that the shift in the variance bounds is statistically significant.

Moreover, we determine whether a candidate kernel x satisfies our variance bounds by testing the null hypothesis

$$\Delta_k^{\mathrm{BK}} \equiv \operatorname{Var}(x) - \operatorname{Var}(q^*) - \delta_k^2 \geq 0. \tag{38}$$

A negative and significant estimate of Δ_k^{BK} indicates that the candidate kernel x fails to satisfy our variance bounds.

Also, we impose the positivity restriction, and test the hypotheses:

$$\tilde{\Delta}^{\mathrm{HJ}} \equiv \operatorname{Var}(x) - \operatorname{Var}(\tilde{q}) \geq 0 \tag{39}$$

$$\tilde{\delta}_k \equiv \frac{\tilde{\lambda}_k - \lambda_{xk}}{\sqrt{\mathrm{E}[(y_k - y_k^*)^2]}} = 0 \tag{40}$$

$$\tilde{\Delta}_k^{\mathrm{BK}} \equiv \operatorname{Var}(x) - \operatorname{Var}(\tilde{q}) - \tilde{\delta}_k^2 \geq 0. \tag{41}$$

The tests above are performed by obtaining estimates of Δ^{HJ}, δ_k, and Δ_k^{BK} (or $\tilde{\Delta}^{\mathrm{HJ}}$, $\tilde{\delta}_k$ and $\tilde{\Delta}_k^{\mathrm{BK}}$), and of the other relevant parameters by exactly identified generalized method of moments (GMM) (see Hansen (1982)). The moment conditions used in the estimation are

$$E[r_a^z(r_a^{z\top}\alpha^z) - \iota^z] = 0 \tag{42}$$

$$E[(r_a^{z\top}\alpha^z - E(r_a^{z\top}\alpha^z))^2] - \mathrm{Var}(q^*) = 0 \tag{43}$$

$$E[r_a^z(y_k - r_a^{z\top}\alpha_{y_k}^z)] = 0 \tag{44}$$

$$E[(y_k - r_a^{z\top}\alpha_{y_k}^z)^2] - E[(y_k - y_k^*)^2] = 0 \tag{45}$$

$$E[-y_k(r_a^{z\top}\alpha_{y_k}^z)] - \lambda_k^* = 0 \tag{46}$$

$$E(-xy_k) - \lambda_{xk} = 0 \tag{47}$$

$$E[(x - E(x))^2 - ((r_a^{z\top}\alpha^z) - E(r_a^{z\top}\alpha^z))^2] - \Delta^{\mathrm{HJ}} = 0 \tag{48}$$

$$E[(x - E(x))^2 - ((r_a^{z\top}\alpha^z) - E(r_a^{z\top}\alpha^z))^2 - \delta_k^2] - \Delta_k^{\mathrm{BK}} = 0. \tag{49}$$

Note that we use the scaled returns r_a^z. When we impose the positivity restriction, similar moment conditions are used, in which the first two are replaced by:

$$E[r_a^z(r_a^{z\top}\tilde{\alpha}^z)^+ - \iota^z] = 0 \tag{50}$$

$$E[((r_a^{z\top}\tilde{\alpha}^z)^+ - E((r_a^{z\top}\alpha^z)^+))^2] - \mathrm{Var}(\tilde{q}) = 0 \tag{51}$$

and the last two moment conditions are replaced by:

$$E[(x - E(x))^2 - ((r_a^{z\top}\tilde{\alpha}^z)^+ - E((r_a^{z\top}\tilde{\alpha}^z)^+))^2] - \tilde{\Delta}^{\mathrm{HJ}} = 0 \tag{52}$$

$$E[(x - E(x))^2 - ((r_a^{z\top}\tilde{\alpha}^z)^+ - E((r_a^{z\top}\tilde{\alpha}^z)^+))^2 - \tilde{\delta}_k^2] - \tilde{\Delta}_k^{\mathrm{BK}} = 0. \tag{53}$$

In order to test the null hypothesis (37), let θ_{δ_k} denote the vector of parameters $[E((y_k^* - y_k)^2), \lambda_k^*, \lambda_{xk}]'$, and let $\hat{\theta}_{\delta_k}$ denote the corresponding GMM estimates. Let us write the quantity δ_k as a function of θ_{δ_k}: $\delta_k = \delta_k(\theta_{\delta_k})$, and let us denote by $\hat{\delta}_k \equiv \delta_k(\hat{\theta}_{\delta_k})$ a consistent estimate of δ_k. By taking a mean-value expansion of $\hat{\delta}_k$ about θ_{δ_k}, and given the asymptotic normality of the GMM estimates $\hat{\theta}_{\delta_k}$, it follows that

$$\sqrt{T}(\hat{\delta}_k - \delta_k) \xrightarrow{L} N(0, \sigma_{\delta_k}^2). \tag{54}$$

Also note that the variance $\sigma_{\delta_k}^2$ is consistently estimated by

$$\hat{\sigma}_{\delta_k}^2 = \left(\left.\frac{\partial \delta_k}{\partial \theta^\top}\right|_{\hat{\theta}_{\delta_k}}\right)\hat{\Sigma}_{\theta_{\delta_k}}\left(\left.\frac{\partial \delta_k}{\partial \theta}\right|_{\hat{\theta}_{\delta_k}}\right) \tag{55}$$

where $\hat{\Sigma}_{\theta_{\delta_k}}$ is a consistent estimate of the covariance matrix of $\hat{\theta}_{\delta_k}$ (see Hansen (1982) and Newey and West (1987)). Under the null hypothesis $\delta_k = 0$, the t-ratio $\hat{\delta}_k/\hat{\sigma}_{\delta_k}$ is asymptotically distributed as a standard normal random variable, and we can construct Wald tests of the null hypothesis (37) with critical values of ± 1.96 and ± 2.58, for two-sided tests at the 5 percent and at the 1 percent levels, respectively. The same arguments can be used to construct tests of the null hypothesis (41).

A test of the null hypothesis (36) can be performed as follows. Let $\hat{\Delta}^{\mathrm{HJ}}$ denote the GMM estimate of Δ^{HJ}, and let $\hat{\sigma}_{\Delta^{\mathrm{HJ}}}$ denote a consistent estimate of its standard error. Under the null that $\Delta^{\mathrm{HJ}} = 0$, the t-ratio $\hat{\Delta}^{\mathrm{HJ}}/\hat{\sigma}_{\Delta^{\mathrm{HJ}}}$ is asymptotically distributed as a standard normal random variable. Since the null hypothesis (36) is an inequality, we can construct Wald tests with critical values of -1.65 and -2.33, for one-sided tests at the 5 percent and at the 1 percent levels, respectively. The same arguments apply to tests of the null hypothesis (38), and of their counterparts when the positivity constraint *is* imposed.

Note that in principle, we could test null hypotheses based on inequalities (26) and (34), where several risk variables are considered *simultaneously*. We do not take this approach in this article for the sake of tractability: It is not possible to graphically represent our variance bounds as a function of the risk premia λ_k when the number of risk variables exceeds two, and the simultaneous test of several inequalities requires the joint estimation of a large number of parameters, which is computationally cumbersome.

IV. Model and Data

This section illustrates our choice of the candidate pricing kernel, asset returns, and risk variables, that we use in the empirical applications.

A. Consumption-Based Pricing Kernels

We consider as a candidate pricing kernel the intertemporal marginal rate of substitution (IMRS) based on the power-utility specification for the cases of time separability and one-period habit persistence. These specifications have been used in empirical work by Grossman and Shiller (1981), Hansen and Singleton (1982), Ferson and Constantinides (1991), and many others. The pricing result deriving from this choice of pricing kernel is often denoted as the consumption-based CAPM (C-CAPM). We do not investigate the case of durability because previous studies, such as Ferson and Constantinides (1991), show that habit persistence has a dominant effect in Euler-equation-based tests, which like ours, use data on consumption of nondurables. Moreover, tests of the H&J regions by Cecchetti, Lam, and Mark (1994) show that for a given rate of relative risk aversion, the assumption of durability leads to stronger rejections of the consumption-based model relative to the time-separable specification.

In what follows, C denotes per-capita real consumption. We assume the log of gross consumption growth, $c_t \equiv \ln(C_t/C_{t-1})$, to be conditionally normal: $c_{t+1} = \mu_{ct} + \varepsilon_{c,t+1}$, where $\varepsilon_{c,t+1} \sim N(0, \sigma^2_{\varepsilon_c t})$. This distributional assumption is common in the literature (see, for example, Cecchetti, Lam, and Mark (1994)) and is roughly consistent with the data set we consider, as argued in the discussion of Section IV.C.

Let ρ denote the time-discount factor, γ the relative rate of risk aversion, and ϕ the habit-persistence parameter. The IMRS between times t and $t+1$ is given by

$$\mathrm{IMRS}_{t+1} \equiv \frac{\rho[(C_{t+1} + \phi C_t)^{-\gamma} + \rho\phi E_{t+1}[(C_{t+2} + \phi C_{t+1})^{-\gamma}]}{(C_t + \phi C_{t-1})^{-\gamma} + \rho\phi E_t[(C_{t+1} + \phi C_t)^{-\gamma}]}, \quad \rho > 0, \quad \gamma > 0, \quad \phi \leq 0. \tag{56}$$

We consider two formulations of the candidate (normalized) pricing kernel based on equation (56) above. In the "conditional" formulation, we have $x_{t+1} \equiv \mathrm{IMRS}_{t+1}/E_t(\mathrm{IMRS}_{t+1})$, and hence $E_t(x_{t+1}) = 1$.

This approach differs from standard tests of the H&J variance bounds, where the IMRS is only required to meet the unconditional moment restriction $\mathrm{E(IMRS)} = \mathrm{E}(1/r_f)$. On the other hand, Gallant, Hansen, and Tauchen (1990) compare the conditional first and second moments of a candidate IMRS to the conditional counterparts of the H&J regions. They find the conditional regions, conditioned at the mean of the data, to be substantially more stringent than the unconditional ones.

In the "unconditional" formulation, we have $x_{t+1} \equiv \mathrm{IMRS}_{t+1}/E(\mathrm{IMRS}_{t+1})$, and hence $\mathrm{E}(x_{t+1}) = 1$, although the conditional expectation of the pricing kernel x, $\mathrm{E}_t(x_{t+1})$, may differ from one.

For tractability, the conditional formulation is implemented only in the time-separable case, $\phi = 0$. Using the definition of c, the IMRS in equation (56) can be written as

$$\mathrm{IMRS}_{t+1} = \frac{\rho e^{-\gamma c_t}[(e^{c_{t+1}}+\phi)^{-\gamma}+\rho\phi e^{-\gamma c_{t+1}}\mathrm{E}_{t+1}((e^{c_{t+2}}+\phi)^{-\gamma})]}{(e^{c_t}+\phi)^{-\gamma}+\rho\phi e^{-\gamma c_{t+1}}\mathrm{E}_t((e^{c_{t+1}}+\phi)^{-\gamma})}. \tag{57}$$

When $\phi = 0$, we can use assumption of lognormality to obtain $\mathrm{IMRS}_{t+1} = \rho \exp[-\gamma(\mu_{ct} + \varepsilon_{c,t+1})]$, which leads to the expression

$$x_{t+1} = \exp(-\gamma\varepsilon_{c,t+1} - \gamma^2\sigma^2_{\varepsilon ct}/2). \tag{58}$$

Note that the specification for x in equation (58) would also obtain for the preferences examined by Abel (1990), when the period utility function is isoelastic and exhibits habit persistence.

When $\phi \neq 0$, closed-form analytical expressions for the expectation $\mathrm{E}_t((e^{c_{t+1}} + \phi)^{-\gamma})$ are not available. Hence, we approximate $(e^{c_{t+1}} + \phi)^{-\gamma}$ around $\mathrm{E}_t(c_{t+1})$ with a third-order polynomial, and we take expectations using the law of motion of c. An alternative approach would be that of discretizing the estimated law of motion of c, and taking expectations based on the transition-probability matrix. This is the route taken in H&J and Cecchetti, Lam, and Mark (1994).

Note that the conditional formulation (58) of the (normalized) pricing kernel x has implications for the degree of relative risk aversion that makes the variance of x consistent with both H&J's and our variance bounds. In fact, assuming $\sigma_{\varepsilon_c t} = \sigma_{\varepsilon_c}$, and using the properties of the lognormal distribution, we have $\mathrm{Var}(x) = \exp(\gamma^2 \sigma_{\varepsilon_c}^2) - 1$, which increases with γ. For comparison, assume that consumption growth is unconditionally lognormal, and $c \sim N(0, \sigma_c^2)$. (This assumption is quite consistent with the data, as argued in the discussion of Section IV.C.) Consider now the "unconditional" normalized kernel $c^{-\gamma}/\mathrm{E}(c^{-\gamma})$ whose variance is given by $\mathrm{Var}[c^{-\gamma}/\mathrm{E}(c^{-\gamma})] = \exp(\gamma^2 \sigma_c^2) - 1$. To the extent that consumption growth is predictable and $\sigma_c > \sigma_{\varepsilon_c}$, the same degree of relative risk aversion γ makes the "conditional" normalized kernel less volatile than its unconditional counterpart. Given a variance lower bound $\bar{v}$, the value of γ that allows the conditional pricing kernel to meet the variance bound is $\gamma = \sqrt{\ln(1 + \bar{v})}/\sigma_{e_c}$. The value of γ that allows the unconditional pricing kernel to meet the variance bound is $\gamma = \sqrt{\ln(1 + \bar{v})}/\sigma_c$. The ratio between the two values is σ_c/σ_{e_c}.

If we estimate the consumption innovations e_c by regressing c on a vector of predictive variables, then $\sigma_c/\sigma_{e_c} = 1/\sqrt{1 - R^2}$, where R^2 is the coefficient of determination for the regression. In other words, as consumption growth becomes more predictable, the higher γ needs to be for the conditional pricing kernel to meet a given variance bound.

Since we consider the conditional formulation of the pricing kernel equation (58), and because of the additional tightness of our variance bounds relative to the H&J bound, we expect our tests to increase the severity of the equity-premium puzzle (see Mehra and Prescott (1985)): Only implausibly high levels of relative risk aversion reconcile stock and bond returns with the behavior of aggregate per-capita consumption growth. In fact, as we increase γ, the risk premia $\lambda_{xk} \equiv -\mathrm{E}(xy_k)$ could also increase in absolute value. To the extent that this increases the quantity $(\lambda_k^* - \lambda_{xk})^2$, it makes inequality (38) more stringent. Values of γ that allow the candidate kernel x to satisfy the null hypothesis $\mathrm{Var}(x) \geq \mathrm{Var}(q^*)$ can still lead to a rejection of the more stringent null equation (38). Thus, the additional variability gained by increasing γ "penalizes" the candidate kernel through its effect on the risk premia λ_{xk}. This intuition makes our analysis similar in spirit to Hansen and Jagannathan (1997), whose measure of misspecification is designed not to reward the variability of a pricing kernel.[7]

B. Asset Returns

We consider monthly-returns data on five stock portfolios and one bond portfolio as in Fama and French (1993).[8] The five stock portfolios are a market portfolio proxy, portfolios of large and small firm stocks, and portfolios of high

[7] Hansen, Heaton, and Luttmer (1995) extend the analysis to the presence of transaction costs and/or short-sale constraints.

[8] We thank Eugene Fama for kindly providing these data.

Table I

Summary Statistics for Asset Returns

We report summary statistics for the asset returns for the period 1963:7–1991:12. All returns are quoted in percentage points per month. Rates of return are as follows: r_{M} is the rate of return on the market portfolio; r_{S} is that for small stocks; and r_{B} is the rate of return on the large stocks; r_{H} is the return on high book-to-market-value stocks; r_{L} is that for the low book-to-market-value stocks; and r_f is the one-month T-bill rate. We deflate all returns by using Consumer Price Index (CPI) inflation. The statistic Corr_τ is the autocorrelation coefficient of order τ.

Panel A: Means, Standard Deviations, and Autocorrelations

Variable	Mean	Std. Dev.	Corr_1	Corr_2	Corr_3	Corr_4	Corr_{12}	Corr_{24}	Corr_{36}
r_{M}	0.5366	4.5804	0.064	−0.036	0.005	−0.013	0.035	−0.027	−0.036
r_f	0.0970	0.3063	0.460	0.423	0.350	0.397	0.417	0.278	0.202
r_{S}	0.8671	5.9940	0.189	−0.025	−0.011	−0.033	0.086	−0.025	0.044
r_{B}	0.5999	4.4439	0.033	−0.037	0.010	−0.025	0.040	−0.021	−0.034
r_{H}	0.9422	5.0066	0.110	−0.044	−0.009	−0.049	0.111	−0.000	0.044
r_{L}	0.5498	5.7120	0.147	−0.018	0.004	−0.034	0.020	−0.041	−0.020

Panel B: Correlation Matrix

Variable	r_{M}	r_f	r_{B}	r_{S}	r_{H}	r_{L}
r_{M}	1.000	0.207	0.896	0.992	0.905	0.961
r_f	0.207	1.000	0.194	0.205	0.201	0.197
r_{B}	0.896	0.194	1.000	0.889	0.957	0.951
r_{S}	0.992	0.205	0.889	1.000	0.928	0.937
r_{H}	0.905	0.201	0.957	0.928	1.000	0.895
r_{L}	0.961	0.197	0.951	0.937	0.895	1.000

and low book-to-market-value stocks. The bond portfolio is the one-month T-bill.

In the following, we briefly describe the portfolio-returns data with reference to the series of Fama and French (1993). The market portfolio returns, r_{M}, correspond to the *RM* series of Fama and French.[9] The small-firm portfolio returns, r_{S}, are the simple average of the returns on the three small-firm portfolios S/L, S/M, and S/H. Similarly, the large-firm portfolio returns, r_{B}, are the simple average of the returns on the three large-firm portfolios B/L, B/M, and B/H. The high book-to-market-value portfolio returns, r_{H}, are the average of the returns on the two high book-to-market-value portfolios S/H and B/H. The low book-to-market-value portfolio returns, r_{L}, are the average of the returns on the two low book-to-market-value portfolios S/L and B/L. The bond portfolio return, r_f, is the one-month T-bill rate, series *TB*, that we use as a proxy for the risk-free rate. All returns are deflated using a CPI inflation series, constructed from the series PZUNEW from CITIBASE.

Table I reports summary statistics for the returns on the stock and bond

[9] For further details on the construction of the stock portfolios, see Fama and French (1993), pp. 8–10.

portfolios, for the period July 1963 to December 1991.

We choose these particular stock portfolio returns because Fama and French (1993) have shown that they capture the relevant risks in the stock market: The regressions of excess stock returns on $r_{\mathrm{M}} - r_f$, the market factor, on $r_{\mathrm{B}} - r_{\mathrm{S}}$, the "size" factor, and on $r_{\mathrm{H}} - r_{\mathrm{L}}$, the "book-to-market" factor, produce intercept estimates that are close to zero.

C. Risk Variables, Consumption Innovations, and Conditioning Variables

We consider five economic variables:

1. *DEF:* The change in the yield on Baa corporate bonds minus the change in the yield on long-term Treasury bonds (annualized yields, percentage points).
2. *TS:* The change in the difference between the yield of a ten-year Treasury note and a three-month T-bill (annualized yields, percentage points).
3. *INF:* The monthly rate of inflation (percentage points per month).
4. *IP:* The monthly growth rate of the industrial production index (percentage points per month).
5. *CG:* The log of the monthly gross growth rate of per-capita real consumption of nondurables and services (percentage points per month).

The yields on Baa corporate bonds, long-term Treasury bonds, ten-year Treasury notes, and three-month T-bills are from the CITIBASE tape, series FYBAAC, FYGL, FYGT10, and FYGM3. The series used to construct consumption data are also from the CITIBASE tape. Monthly real consumption of nondurable goods and services are the GMCN and GMCS series deflated by the corresponding deflator series GMDN and GMDS. Per capita quantities are obtained by using data on resident population, series POPRES.

Summary statistics for the five economic variables for the June 1963 to November 1991 period are reported in Table II. Note that we use this sample period, rather than July 1963 to December 1991, because the five economic variables are later used as instruments to scale returns, and thus lagged one period. Therefore, the statistics reported in the table are those of the instruments z. We also test whether the skewness and kurtosis of log of consumption growth is consistent with the assumption of normality. We obtain a skewness of 0.0771 and a kurtosis of 0.3034, which do not lead to a rejection of the normality assumption at conventional significance levels.

We have chosen these economic variables on the basis of earlier studies which find that they command nonzero risk premia in empirical investigations of multiple-beta and of multiple-factor models. Similar variables have been used, for example, in Chen, Roll, and Ross (1986), Burmeister and McElroy (1988), and Ferson and Harvey (1991).

We assume the risk variables y_k to have mean zero and unit variance, and that they are conditionally orthogonal to each other. We also assume that the innovations in the log of consumption growth $\varepsilon_{c,t+1}$ have mean zero, and are

Table II

Summary Statistics for Economic Variables

We report summary statistics for the economic variables for the period 1963:6–1991:11. *DEF* is the change in the yield on Baa corporate bonds minus the change in the yield on long-term Treasury bonds (annualized yields, percentage points). *TS* is the change in the difference between the yield of a 10-year Treasury note and a 3-month T-bill (annualized yields, percentage points). *INF* is the monthly rate of inflation (percentage points per month). *IP* is the monthly growth rate of the industrial production index (percentage points per month). *CG* is the log of the monthly gross growth rate of per capita real consumption of nondurables and services (percentage points per month). The statistic Corr_τ is the autocorrelation coefficient of order τ.

Panel A: Means, Standard Deviations, and Autocorrelations

Variable	Mean	Std. Dev.	Corr_1	Corr_2	Corr_3	Corr_4	Corr_{12}	Corr_{24}	Corr_{36}
DEF	0.0022	0.1728	0.181	−0.157	−0.052	0.035	−0.026	0.017	0.027
TS	0.0054	0.4484	0.232	−0.156	−0.041	−0.076	−0.153	−0.004	−0.015
INF	0.4425	0.3415	0.572	0.535	0.439	0.442	0.430	0.177	0.108
IP	0.2661	0.8525	0.374	0.266	0.219	0.163	−0.014	−0.250	−0.019
CG	0.1646	0.4016	−0.273	0.118	0.146	−0.032	−0.047	−0.178	−0.021

Panel B: Correlation Matrix

Variable	*DEF*	*TS*	*INF*	*IP*	*CG*
DEF	1.000	−0.081	0.072	−0.150	0.008
TS	−0.081	1.000	0.040	−0.215	−0.119
INF	0.072	0.040	1.000	−0.094	−0.278
IP	−0.150	−0.215	−0.094	1.000	0.200
CG	0.008	−0.119	−0.278	0.200	1.000

conditionally orthogonal to the information set at time t. To generate a set of variables that satisfies these conditions, we estimate a vector autoregression for the five economic variables. We then use the orthogonalized residuals as the risk variables y_k. The residuals from the consumption-growth equation are used as the innovations e_c, and drive the process for the log of consumption growth. We report results of the estimation and summary statistics for the residuals in Table III.

Panel A, Table III, shows that in a two-tailed test at the 5 percent significance level, each variable is significantly explained by its own lag. *DEF* is significantly predicted by *TS* and *INF*; *TS* is significantly predicted by *DEF*; *INF* is significantly predicted by *CG*; *IP* is significantly predicted by *DEF*; and *CG* is significantly predicted by *INF* and *IP*.

Panel B, Table III, shows that the larger correlations among residuals of the VAR are between innovations in consumption growth and innovations in inflation and industrial production growth.

We assume homoskedasticity of the VAR residuals, and set σ_{ct} equal to the standard deviation of the residual of the consumption-growth equation. This residual exhibits a skewness of 0.1877 and a kurtosis of 0.6178; thus, the series that we use as innovations in the log of consumption growth is roughly

Table III

Summary Statistics for VAR Estimation and Residuals

We report vector autoregression (VAR) parameter estimates and diagnostics and summary statistics for the VAR residuals, for the period 1963:7–1991:12. Each row in Panel A corresponds to an equation of the VAR. *DEF* is the change in the yield on Baa corporate bonds minus the change in the yield on long-term Treasury bonds (annualized yields, percentage points). *TS* is the change in the difference between the yield of a 10-year Treasury note and a 3-month T-bill (annualized yields, percentage points). *INF* is the monthly rate of inflation (percentage points per month). *IP* is the monthly growth rate of the industrial production index (percentage points per month). *CG* is the log of the monthly gross growth rate of per capita real consumption of nondurables and services (percentage points per month). The t-ratios of the VAR estimates are reported in parentheses.

Panel A: VAR Parameter Estimates

Equation	*DEF*	*TS*	*INF*	*IP*	*CG*	$\bar{R}^2$
DEF	0.1502	−0.0545	0.0709	−0.0139	−0.0338	0.0704
	(2.817)	(−2.618)	(2.569)	(−1.241)	(−1.413)	
TS	0.4937	0.2279	0.0318	−0.0533	0.0263	0.0940
	(3.614)	(4.275)	(0.450)	(−1.860)	(0.429)	
INF	0.0083	0.0060	0.5977	−0.0111	0.0842	0.3269
	(0.092)	(0.171)	(12.860)	(−0.589)	(2.093)	
IP	−1.2194	−0.1538	−0.2173	0.2984	0.1497	0.2079
	(−5.016)	(−1.621)	(−1.728)	(5.848)	(1.376)	
CG	−0.1793	−0.0225	−0.2343	0.0718	−0.3613	0.1371
	(−1.503)	(−0.483)	(−3.795)	(2.869)	(−6.763)	

Panel B: Residuals Correlation Matrix

Variable	*DEF*	*TS*	*INF*	*IP*	*CG*
DEF	1.000	−0.108	−0.031	−0.078	0.030
TS	−0.108	1.000	0.006	−0.109	−0.083
INF	−0.031	0.006	1.000	−0.002	−0.230
IP	−0.078	−0.109	−0.002	1.000	0.178
CG	0.030	−0.083	−0.230	0.178	1.000

consistent with the assumption of normality (the corresponding p-values are 0.158 and 0.021, respectively).

The residuals from all equations are then made orthogonal and normalized by their standard deviation, using a Choleski decomposition of their covariance matrix. The order of the decomposition is *DEF*, *TS*, *INF*, *IP*, and *CG*. The transformed residuals are denoted y_{DEF}, y_{TS}, y_{INF}, y_{IP}, and y_{CG}. In light of the evidence from Panel B, Table III, the orthogonalization especially affects the residuals from the consumption-growth equation.

Finally, the set of instruments to obtain the scaled returns r_a^z is chosen to be the vector

$$z = [1, DEF, TS, INF, IP, CG]^\top. \tag{59}$$

Table IV

Region-Subset Tests

We test the null hypothesis H_0 that the indicated set of asset returns contains all the relevant information concerning the minimum-variance admissible (normalized) kernel q^*. The alternative hypothesis H_1 is that a larger set of asset returns is needed to summarize the relevant information concerning q^*. The asset returns considered are: r_M, the rate of return on the market portfolio; r_f, the one-month T-bill rate; r_S and r_B, the rates of return on small stocks and large stocks, respectively; $r_{BS} \equiv r_f + r_B - r_S$; r_H and r_L, the rates of return on high book-to-market-value and low book-to-market-value stocks, respectively; and $r_{HL} \equiv r_f + r_H - r_L$. (All returns are deflated using Consumer Price Index (CPI) inflation.) We report the J-statistic, which is asymptotically distributed as a chi-square random variable with degrees of freedom equal to the number of overidentifying restrictions (see Hansen 1982). The tests are performed with ("Positivity Restriction") and without ("Unrestricted") the positivity constraint. The tests assume errors to be serially uncorrelated (SU), or to exhibit serial correlation of moving average form up to 12 lags (MA). The J-statistic is robust to conditional heteroskedasticity of the errors. In case of singularity of the covariance matrix of the errors (in the MA case), we use the Newey and West (1987) correction.

			Unrestricted		Positivity Restriction	
H_0	H_1	df	SU	MA	SU	MA
r_M, r_f	All returns	4	16.25	5.37	16.25	5.37
r_M, r_f, r_B, r_S	All returns	2	5.85	2.30	5.83	2.27
r_M, r_f, r_{BS}, r_{HL}	All returns	2	0.52	0.33	0.46	0.30
r_M, r_f, r_{BS}, r_{HL}	All returns, scaled	20	39.96	15.01	40.07	15.11

D. Region-Subset Tests

In order to identify a set of assets and instruments that summarizes the properties of the pricing kernel, we perform some region-subset tests. We follow a procedure described in Hansen, Heaton, and Luttmer (1995).[10] Specifically, we test zero restrictions on the coefficient vectors α^z and $\tilde{\alpha}^z$ that relate q^* and $\tilde{q}$ to r_a^z.

The tests are performed as follows: The coefficient vectors α^z and $\tilde{\alpha}^z$ must satisfy the moment conditions $\mathrm{E}[r_a^z(r_a^{z\top}\alpha^z) - \iota^z] = 0$ and $\mathrm{E}[r_a^z(r_a^{z\top}\tilde{\alpha}^z)^+ - \iota^z] = 0$, respectively. When the elements of α^z and $\tilde{\alpha}^z$ are unrestricted, the two coefficient vectors can be estimated by exactly identified GMM: The number of moment conditions equals the number of parameters. When some of the elements of α^z and $\tilde{\alpha}^z$ are restricted to zero, then the remaining elements can be estimated by an overidentified GMM in which the number of overidentifying restrictions equals the number of the restricted elements of the coefficient vectors. The J-statistic produced by the GMM estimation (see Hansen (1982)) is asymptotically distributed as a chi-square random variable with degrees of freedom equal to the number of overidentifying restrictions.

The results of the region-subset tests are reported in Table IV.

[10] Similar tests are performed in Snow (1991).

We test the following four hypotheses:

1. r_f and $r_{\rm M}$ "price" all returns ($r_{\rm M}$, r_f, $r_{\rm B}$, $r_{\rm S}$, $r_{\rm H}$, and $r_{\rm L}$).
2. r_f, $r_{\rm M}$, $r_{\rm B}$, and $r_{\rm S}$ price all returns.
3. r_f, $r_{\rm M}$, $r_{\rm BS} \equiv r_f + r_{\rm B} - r_{\rm S}$, and $r_{\rm HL} \equiv r_f + r_{\rm H} - r_{\rm L}$ price all returns.
4. r_f, $r_{\rm M}$, $r_{\rm BS}$, and $r_{\rm HL}$ price all returns, *scaled* by the conditioning variables.

In light of the results of Snow (1991) and of Fama and French (1993), we would expect the first two hypotheses to be rejected, but not the third. Also, beginning with Hansen and Singleton (1982), many studies have established the relevance of conditioning information in testing asset-pricing models on post-war U.S. data, and we would expect the fourth hypothesis to be rejected. The four hypotheses are tested with and without the positivity constraint. We calculate the covariance matrix of the estimates by assuming that the pricing errors are either serially uncorrelated (SU), or exhibit serial correlation of moving-average form up to twelve lags (MA) (see Hansen (1982)). When the estimated covariance matrix cannot be inverted, we use the Newey and West (1987) correction.

While the positivity restriction makes little if any difference, the test results are quite sensitive to the assumed error structure. This feature is common to other studies of the H&J regions. For example, see De Santis (1995) and Downs and Snow (1994). The sensitivity of the magnitude of the test statistics to the assumed error structure is explained by the serial dependence in real T-bill returns, r_f, as shown in Table I, Panel A.

When we assume that errors are serially uncorrelated, we reject the null hypothesis in the first three tests. When we assume that errors have a moving-average structure, we do not reject the null in any of the four tests. This evidence indicates that the three factor portfolios of Fama and French (1993) are needed to capture the systematic variation of stock returns. The remaining empirical analysis is performed using the four asset returns $r_{\rm M}$, r_f, $r_{\rm BS}$, and $r_{\rm HL}$, scaled by the conditioning variables.

V. Test Results

Here, we test whether the candidate consumption-based pricing kernels, introduced in Section IV, meet the H&J variance bound; whether our bounds are strictly more stringent than the H&J bound; and whether the consumption-based pricing kernels meet our variance bounds.

A. A First Set of Tests

Following the methodology discussed in Section III, we obtain estimates of $\text{Var}(q^*)$ and $\text{Var}(\tilde{q})$, which equal 0.1164 and 0.1171. The corresponding t-ratios are 2.679 and 2.633, respectively. In the discussion of the test results, we focus on the standard errors calculated under the assumption of serial correlation in the GMM errors, the MA case. For sake of comparison, the tables also report standard errors calculated under the no-serial-correlation assumption, the SU case.

In the first set of tests, presented in Table V, we focus on the conditional consumption-based candidate pricing kernel with time separability, specified in equation (58). We set γ equal to the value for which x exactly satisfies the H&J variance bound ($\Delta^{HJ} = 0$; and $\tilde{\Delta}^{HJ} = 0$, with positivity restriction) for the data set under scrutiny. This parameter choice highlights the incremental information in our variance bounds relative to the H&J bound. Without the positivity restriction, $\gamma = 91.289$. With the positivity restriction, $\gamma = 91.526$. Note that these are not intended to be estimates of γ, but reference values of the relative rate of risk aversion. Also note that the normalized candidate pricing kernel in equation (58) does not depend on ρ. Therefore, we do not need to choose a value for this time-discount parameter.

Table V reports estimates of $E[(y_k - y_k^*)^2]$, λ_k^*, λ_{xk}, δ_k, and Δ_k^{BK} for the five risk variables. We also report estimates of $\tilde{\lambda}_k$, λ_{xk}, $\tilde{\delta}_k$ and $\tilde{\Delta}_k^{BK}$, when the positivity restriction is imposed.

The first column in Table V measures how closely the mimicking portfolios can replicate the various risk variables. All estimates are precise, and indicate that innovations in inflation are the most closely tracked by a mimicking portfolio: The residual variance $E[(y_{INF} - y_{INF}^*)^2]$ does not exceed 34 percent. The other four variables are not as well tracked, with an unexplained variance that exceeds 91 percent in the case of y_{IP}.

Note that the residual variances $E[(y_k^* - y_k)^2]$ play a crucial role in placing bounds on the risk premia (see equations (27) and (35)). In fact, the size of the interval within which a risk premium needs to lie is given by $2\sqrt{[\mathrm{Var}(q) - \mathrm{Var}(q^*)]E[(y_k - y_k^*)^2]}$. Thus, the derivative of this interval size with respect to $\sqrt{[\mathrm{Var}(q) - \mathrm{Var}(q^*)]}$ equals $2\sqrt{E[(y_k - y_k^*)^2]}$. In the case of inflation risk, a unit increase in $\sqrt{[\mathrm{Var}(q) - \mathrm{Var}(q^*)]}$ translates into a 1.15 increase in the size of the risk-premium interval. In the case of industrial production risk, a unit increase in $\sqrt{[\mathrm{Var}(q) - \mathrm{Var}(q^*)]}$ translates into a 1.92 increase in the size of the risk-premium interval.

The second column in Table V reports the risk premia assigned by the minimum-variance normalized pricing kernel q^*. The variables y_{DEF}, y_{TS}, y_{INF}, and y_{CG} all command positive risk premia. However, the premium on the risk variable y_{IP} is negative, although smaller (in absolute value) than the premia on the other variables.

Both Chen, Roll, and Ross (1986) and Ferson and Harvey (1991) also find that the credit spread and consumption growth command positive risk premia. Ferson and Harvey find that changes in the slope of the term structure command a positive premium. The negative premium on y_{IP} is consistent with the results of Burmeister and McElroy (1988), who find that a similar business-cycle indicator—innovations in the growth rate of real final sales—commands a negative risk premium.

The magnitude and statistical significance of these risk premia can be better understood by recalling that they represent estimates of the mean cash flows generated by the portfolios that best mimic the risk variables, financed at the

Table V

Variance Bounds Tests for Conditional Case, Time Separability, Fixed γ

We report results of tests of the conditional (normalized) pricing kernel

$$x_{t+1} = \mathrm{IMRS}_{t+1}/\mathrm{E}_t(\mathrm{IMRS}_{t+1})$$

where IMRS_{t+1} denotes the intertemporal marginal rate of substitution between t and $t + 1$, for the time-separable case. We have

$$x_{t+1} = \exp(-\gamma\varepsilon_{c,t+1} - \gamma^2\sigma^2_{\varepsilon_c}/2)$$

where ε_c is the innovation in the log of consumption growth, c; σ_{ε_c} is the conditional volatility of c; and γ is the relative risk aversion parameter. y_{DEF}, y_{TS}, y_{INF}, y_{IP}, and y_{CG} are the residuals from a vector autoregression. *DEF* is the change in the yield on Baa corporate bonds minus the change in the yield on long-term Treasury bonds (annualized yields, percentage points). *TS* is the change in the difference between the yield of a 10-year Treasury note and a 3-month T-bill (annualized yields, percentage points). *INF* is the monthly rate of inflation (percentage points per month). *IP* is the monthly growth rate of the industrial production index (percentage points per month). *CG* is the log of the monthly gross growth rate of per capita real consumption of nondurables and services (percentage points per month). These are orthogonal and normalized by their standard deviation by means of a Choleski decomposition of their covariance matrix (the order of the decomposition is *DEF*, *TS*, *INF*, *IP*, and *CG*). λ^*_k is the risk premium assigned by the minimum-variance kernel to the risk variable y_k. λ_{xk} is the risk premium assigned by the candidate kernel x to the risk variable y_k. δ_k is the difference between our variance bound and the Hansen and Jagannathan (1991) variance bound. Δ_k is the difference between the variance of the candidate kernel x and our variance bound. $\tilde{\lambda}_k$ is the risk premium assigned by the nonnegative minimum-variance kernel to the risk variable y_k. $\tilde{\delta}_k$ is the difference between our variance bound and the Hansen and Jagannathan (1991) variance bound when the positivity restriction is imposed. $\tilde{\Delta}_k$ is the difference between the variance of the candidate kernel x and our variance bound when the positivity restriction is imposed. The relative risk aversion parameter is set equal to $\gamma = 91.289$ ($\Delta^{\mathrm{HJ}} = 0$) and $\gamma = 91.526$ ($\tilde{\Delta}^{\mathrm{HJ}} = 0$). The tests are performed with and without the positivity constraint. The t-ratios, reported in parentheses, are robust to conditional heteroskedasticity. We assume errors to be serially uncorrelated (SU), or to exhibit serial correlation of moving average-form up to 12 lags (MA). In case of singularity of the covariance matrix of the errors (in the MA case), we use the Newey and West (1987) correction.

		Unrestricted				Positivity Restriction			
Variable	$\mathrm{E}[(y_k - y^*_k)^2]$	λ^*_k	λ_{xk}	δ_k	Δ_k	$\tilde{\lambda}_k$	λ_{xk}	$\tilde{\delta}_k$	$\tilde{\Delta}_k$
y_{DEF}	0.8471	0.0384	0.0091	0.0318	−0.0010	0.0383	0.0091	0.0316	−0.0010
SU	(10.338)	(0.751)	(0.152)	(0.842)	(−0.025)	(0.747)	(0.152)	(0.833)	(−0.024)
MA	(6.747)	(0.836)	(0.175)	(0.964)	(−0.022)	(0.837)	(0.175)	(0.963)	(−0.021)
y_{TS}	0.7413	0.0293	−0.0300	0.0689	−0.0047	0.0287	−0.0301	0.0683	−0.0047
SU	(6.535)	(0.635)	(−0.516)	(1.597)	(−0.116)	(0.623)	(−0.518)	(1.582)	(−0.112)
MA	(3.845)	(0.602)	(−0.711)	(1.678)	(−0.103)	(0.590)	(−0.712)	(1.650)	(−0.099)
y_{INF}	0.3327	0.0176	−0.0876	0.1824	−0.0333	0.0183	−0.0879	0.1840	−0.0339
SU	(11.504)	(0.321)	(−1.351)	(3.378)	(−0.775)	(0.334)	(−1.354)	(3.411)	(−0.772)
MA	(5.057)	(0.243)	(−1.058)	(2.797)	(−0.680)	(0.252)	(−1.060)	(2.831)	(−0.675)
y_{IP}	0.9193	−0.0041	0.0615	−0.0685	−0.0047	−0.0047	0.0617	−0.0693	−0.0048
SU	(9.545)	(−0.072)	(1.001)	(−2.229)	(−0.119)	(−0.083)	(1.003)	(−2.251)	(−0.119)
MA	(8.751)	(−0.074)	(1.002)	(−1.892)	(−0.105)	(−0.085)	(1.004)	(−1.915)	(−0.105)
y_{CG}	0.8885	0.0314	0.3101	−0.2957	−0.0874	0.0303	0.3108	−0.2976	−0.0886
SU	(11.431)	(0.595)	(4.973)	(−8.827)	(−2.335)	(0.576)	(4.981)	(−8.947)	(−2.311)
MA	(12.052)	(0.552)	(4.752)	(−8.115)	(−2.207)	(0.532)	(4.761)	(−8.275)	(−2.179)

riskless rate. The unconditional Sharpe ratios of the cash flows generated by these positions are given by $\lambda_k^*/\sqrt{\mathrm{Var}(y_k^*)} = \lambda_k^*/\sqrt{1 - \mathrm{E}[(y_k - y_k^*)^2]}$. For example, the monthly Sharpe ratio of the portfolio mimicking y_{CG} equals 0.0939, with a t-ratio of 0.523. By comparison, the Sharpe ratio of the market portfolio equals 0.0961, with a t-ratio of 1.877. (These standard errors are calculated following a procedure similar to that used for the estimates of δ_k and $\tilde{\delta}_k$.) Hence, the Sharpe ratio of the portfolio that best mimics consumption risk has a magnitude comparable to that of the market portfolio, while the lower significance can be attributed to the sampling variability around the estimates of the portfolio quantities α_{yk}^z.

The estimates of the risk premia λ_k^* can also be interpreted in light of the analysis of Section I. Specifically, when an economic variable cannot be exactly replicated by the returns on the available assets, then the estimate of its risk premium is the result of a joint hypothesis, one part of which concerns the functional form of the pricing kernel and/or of the return-generation process, and the other concerning the relevant set of asset returns.

As we illustrate in Section II, the risk premia λ_k^* are those implied by choosing the minimum-variance (normalized) kernel q^*, and by choosing the set of asset returns r. Hence, the estimates of the risk premia λ_k^* are a useful benchmark against which we can compare the results of the studies of multiple-beta and multiple-factor models. These models assume pricing kernels and/or return-generation processes that are linear in the risk variables and consider different sets of asset returns.[11]

The third column in Table V reports the risk premia assigned by the candidate pricing kernel x. The premia differ both in sign and size from those assigned by q^*. Specifically, the variables y_{DEF}, y_{IP}, and y_{CG} command positive risk premia, while y_{TS} and y_{INF} command negative risk premia. The difference in the size of the risk premia is most apparent for the risk variable y_{CG}. The estimates of λ_{CG}^* and λ_{xCG} differ by one order of magnitude. A discrepancy between the risk premia λ_{xk} implied by an explicit asset-pricing model and the premia λ_k^* is not per se an indication of the failure of the model. A candidate (normalized) pricing kernel x could price correctly the assets r under consideration, i.e., $-\mathrm{E}(xy_k^*) = \lambda_k^*$, but it could covary with risks that are orthogonal to the (augmented) asset-return space, and it might assign risk premia λ_{xk} to the risk variables y_k that differ from λ_k^*.

For our variance bounds to be more stringent than those of H&J, the risk premia assigned by q^* and x must be very different. This is the case of consumption risk. The difference in risk premia is then amplified if the variance of the risk variable that is not explained by asset returns is small. This is the case of inflation risk. Consequently, both inflation and consumption risk imply variance bounds that are substantially more stringent than the H&J bound. In fact, for these two risk variables, the estimates of the shifts in the variance bounds δ_k reported in Table V are statistically significant in a two-

[11] For example, see the previously cited articles by Chen, Roll, and Ross (1986), Ferson and Harvey (1991), Fama and French (1993), and Burmeister and McElroy (1988).

Table VI

Variance Bounds Tests for Conditional Case, Time Separability, Changes in γ

We report results of tests of the conditional (normalized) pricing kernel

$$x_{t+1} = \mathrm{IMRS}_{t+1}/\mathrm{E}_t(\mathrm{IMRS}_{t+1})$$

where IMRS_{t+1} denotes the intertemporal marginal rate of substitution between t and $t+1$, for the time-separable case. We have

$$x_{t+1} = \exp(-\gamma\varepsilon_{c,t+1} - \gamma^2\sigma_{\varepsilon c}^2/2)$$

where ε_c is the innovation in the log of consumption growth, c; σ_{ε_c} is the conditional volatility of c; and γ is the relative risk aversion parameter. $\tilde{\Delta}^{\mathrm{HJ}}$ is the difference between the variance of the candidate kernel x and the Hansen and Jagannathan (1991) bound when the positivity restriction is imposed. $\tilde{\delta}_{INF}$ and $\tilde{\delta}_{CG}$ are the differences between our variance bound and the Hansen and Jagannathan (1991) variance bound, when the positivity restriction is imposed, for inflation and consumption risk, respectively. $\tilde{\Delta}_{INF}^{\mathrm{BK}}$ and $\tilde{\Delta}_{CG}^{\mathrm{BK}}$ are the differences between the variance of the candidate kernel x and our variance bound, when the positivity restriction is imposed, for inflation and consumption risk, respectively. The t-ratios, reported in parentheses, are robust to conditional heteroskedasticity. We assume errors to be serially uncorrelated (SU) or to exhibit serial correlation of moving average-form up to 12 lags (MA). In case of singularity of the covariance matrix of the errors (in the MA case), we use the Newey and West (1987) correction.

γ	$\tilde{\Delta}^{\mathrm{HJ}}$	$\tilde{\delta}_{INF}$	$\tilde{\Delta}_{INF}^{\mathrm{BK}}$	$\tilde{\delta}_{CG}$	$\tilde{\Delta}_{CG}^{\mathrm{BK}}$
40	−0.955	−0.0938	−0.1043	0.1146	−0.1086
SU	(−2.521)	(−2.651)	(−2.645)	(4.457)	(−2.902)
MA	(−2.142)	(−2.270)	(−2.258)	(3.865)	(−2.572)
60	−0.0681	−0.1274	−0.0843	0.1863	−0.1028
SU	(−1.785)	(−3.109)	(−2.092)	(6.745)	(−2.739)
MA	(−1.519)	(−2.627)	(−1.796)	(5.972)	(−2.490)
80	−0.0287	−0.1628	−0.0552	0.2571	−0.0948
SU	(−0.733)	(−3.347)	(−1.313)	(8.355)	(−2.499)
MA	(−0.629)	(−2.791)	(−1.138)	(7.597)	(−2.327)
100	0.0242	−0.2000	−0.0158	0.3272	−0.0289
SU	(0.547)	(−3.435)	(−0.345)	(9.235)	(−2.143)
MA	(0.504)	(−2.844)	(−0.305)	(8.648)	(−2.039)
120	0.0927	−0.2388	0.0357	0.3966	−0.0646
SU	(1.869)	(−3.430)	(0.663)	(9.484)	(−1.613)
MA	(1.725)	(−2.840)	(0.602)	(9.132)	(−1.569)

sided test at the 1 percent level: The t-ratios of the estimates of δ_{INF} and δ_{CG} exceed 2.58.

This additional tightness can also be appreciated using two other "metrics." One metric is the relative increase in the minimum volatility of a (normalized) pricing kernel:

$$\sqrt{\mathrm{Var}(q^*) + \delta_k^2}\Big/\sqrt{\mathrm{Var}(q^*)} - 1. \tag{60}$$

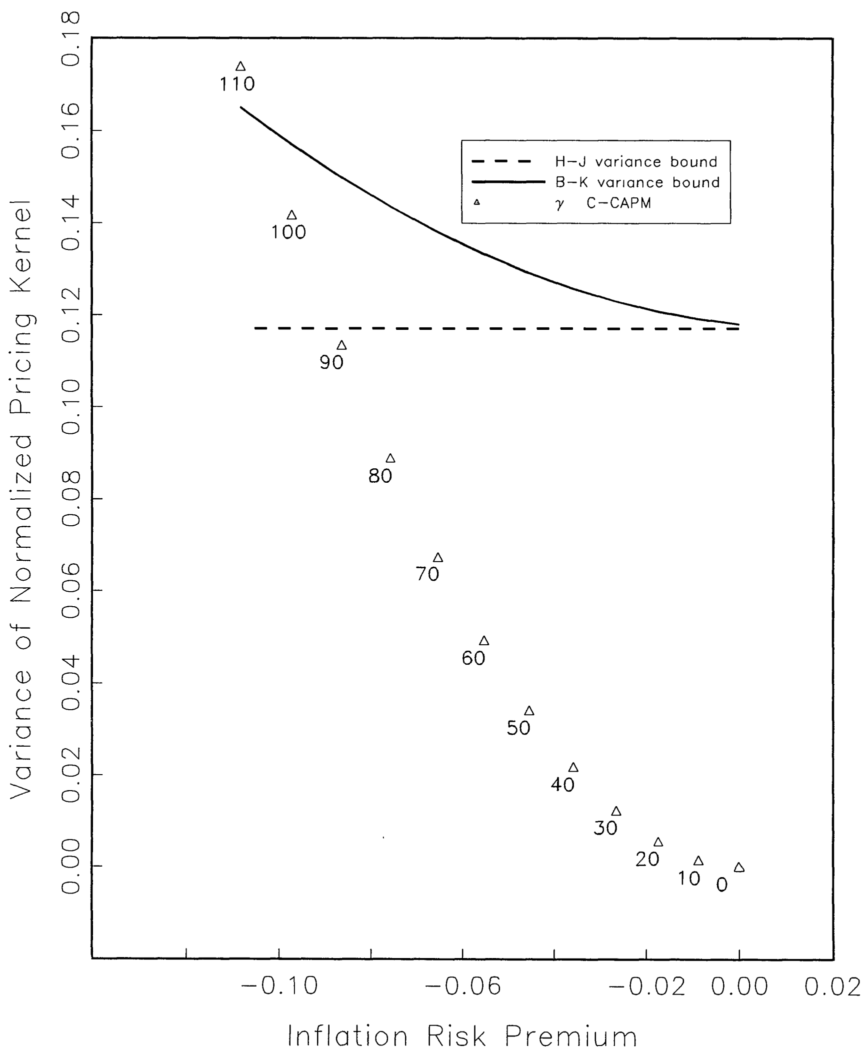

Figure 1. Tests of the Variance Bounds, Inflation. Figure 1 plots the Hansen and Jagannathan (1991) variance bound (H–J) as a flat line, and our variance bounds (B–K) as a parabola, in the risk premium-variance space. We impose the positivity restriction on the pricing kernel. The figure also plots the locus of risk premium-variance pairs implied by the conditional candidate consumption-based kernel (C-CAPM),

$$x_{t+1} = \text{IMRS}_{t+1}/\text{E}_t(\text{IMRS}_{t+1})$$

where IMRS_{t+1} denotes the intertemporal marginal rate of substitution between t and $t + 1$, for the time-separable case.

The increase is 13.41 percent in the case of inflation risk and 32.32 percent in the case of consumption risk. The second metric is defined by the increase in the coefficient of relative risk aversion γ needed for the consumption-based

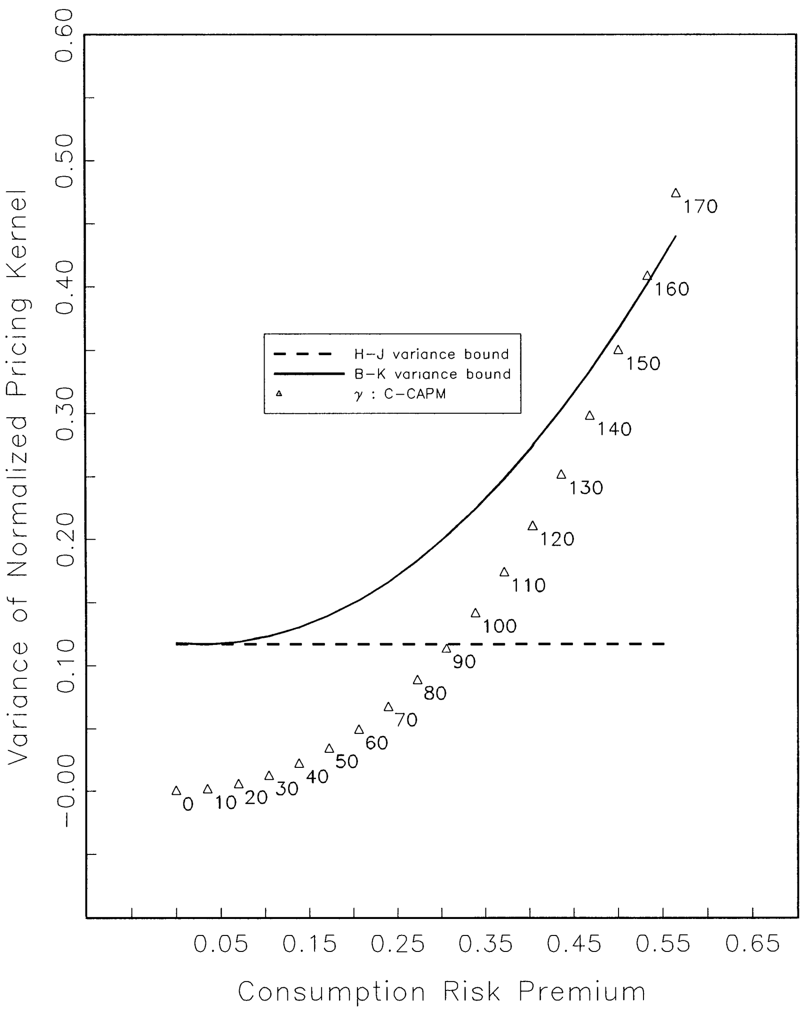

Figure 2. Tests of the Variance Bounds, Consumption. Figure 2 plots the Hansen and Jagannathan (1991) variance bound (H–J) as a flat line, and our variance bounds (B–K) as a parabola, in the risk premium-variance space. We impose the positivity restriction on the pricing kernel. The figure also plots the locus of risk premium-variance pairs implied by the conditional candidate consumption-based kernel (C-CAPM),

$$x_{t+1} = \mathrm{IMRS}_{t+1}/\mathrm{E}_t(\mathrm{IMRS}_{t+1})$$

where IMRS_{t+1} denotes the intertemporal marginal rate of substitution between t and $t + 1$, for the time-separable case.

kernel to meet the variance bounds. The smallest values of γ that satisfy our variance bounds are $\gamma = 106.242$ for the risk variable y_{INF} and $\gamma = 157.277$ for the risk variable y_{CG}, which represent increases of 16.38 percent and 71.28 percent, respectively, in the relative rate of risk aversion.

Table VII

Variance Bounds Tests for Unconditional Case, Changes in γ

We report results of tests of the unconditional (normalized) pricing kernel

$$x_{t+1} = \mathrm{IMRS}_{t+1}/\mathrm{E}(\mathrm{IMRS}_{t+1})$$

where

$$\mathrm{IMRS}_{t+1} \equiv \frac{\rho[(C_{t+1} + \phi C_t)^{-\gamma} + \rho\phi \mathrm{E}_{t+1}[(C_{t+2} + \phi C_{t+1})^{-\gamma}]}{(C_t + \phi C_{t-1})^{-\gamma} + \rho\phi \mathrm{E}_t[(C_{t+1} + \phi C_t)^{-\gamma}]}, \qquad \rho > 0, \quad \gamma > 0, \quad \phi \leq 0.$$

ρ is the time-discount parameter, C is consumption, ϕ is the habit-persistence parameter, and γ is the relative risk aversion coefficient. We consider the time-separable case, $\phi = 0$ (Panel A), and the case of habit persistence, $\phi = -0.5$ (Panel B). ρ is set equal to 1. $\tilde{\Delta}^{\mathrm{HJ}}$ is the difference between the variance of the candidate kernel x and the Hansen and Jagannathan (1991) bound when the positivity restriction is imposed. $\tilde{\delta}_{INF}$ and $\tilde{\delta}_{CG}$ are the differences between our variance bound and the Hansen and Jagannathan (1991) variance bound when the positivity restriction is imposed, for inflation and consumption risk, respectively. $\tilde{\Delta}^{\mathrm{BK}}_{INF}$ and $\tilde{\Delta}^{\mathrm{BK}}_{CG}$ are the differences between the variance of the candidate kernel x and our variance bound when the positivity restriction is imposed, for inflation and consumption risk, respectively. The t-ratios, reported in parentheses, are robust to conditional heteroskedasticity. We assume errors to be serially uncorrelated (SU), or to exhibit serial correlation of moving average-form up to 12 lags (MA). In case of singularity of the covariance matrix of the errors (in the MA case), we use the Newey and West (1987) correction.

Panel A: Tests, Time Separability					
γ	$\tilde{\Delta}^{\mathrm{HJ}}$	$\tilde{\delta}_{INF}$	$\tilde{\Delta}^{\mathrm{BK}}_{INF}$	$\tilde{\delta}_{CG}$	$\tilde{\Delta}^{\mathrm{BK}}_{CG}$
30	−0.1015	−0.0783	−0.1076	0.0937	−0.1103
SU	(−2.692)	(−2.334)	(−2.765)	(3.487)	(−2.941)
MA	(−2.266)	(−1.934)	(−2.335)	(2.874)	(−2.558)
50	−0.0753	−0.1123	−0.0879	0.1720	−0.1049
SU	(−1.979)	(−2.880)	(−2.212)	(5.926)	(−2.773)
MA	(−1.667)	(−2.328)	(−1.881)	(4.860)	(−2.475)
70	−0.034	−0.1486	−0.0569	0.2502	−0.0974
SU	(0.890)	(−3.157)	(−1.377)	(7.660)	(−2.525)
MA	(−0.753)	(−2.499)	(−1.185)	(6.332)	(−2.301)
90	0.0218	−0.1873	−0.0133	0.3283	−0.0860
SU	(0.518)	(−3.255)	(−0.299)	(8.636)	(−2.165)
MA	(0.445)	(−2.551)	(−0.262)	(7.276)	(−2.000)
110	0.0969	−0.2285	0.0447	0.4064	−0.0683
SU	(1.968)	(−3.250)	(0.867)	(8.966)	(−1.650)
MA	(1.763)	(−2.549)	(0.783)	(7.754)	(−1.534)

The fifth column of Table V reports estimates and t-ratios for Δ^{BK}_k for the five risk variables. In the case of consumption risk, the t-ratio of the estimate of Δ^{BK}_k exceeds two (in absolute value). This implies that even a value of γ as high as 90 leads to a strong rejection of the standard consumption-based kernel.

Table V also reports the results of our tests when the positivity restriction is imposed. The results are similar to those discussed above, and confirm our conclusions.

Table VII—*Continued*

Panel B: Tests, Habit Persistence					
γ	$\tilde{\Delta}^{HJ}$	$\tilde{\delta}_{INF}$	$\tilde{\Delta}_{INF}^{BK}$	$\tilde{\delta}_{CG}$	$\tilde{\Delta}_{CG}^{BK}$
2	−0.1138	−0.0484	−0.1162	0.0157	−0.1141
SU	(−3.025)	(−1.557)	(−3.027)	(0.605)	(−3.037)
MA	(−2.546)	(−1.291)	(−2.547)	(0.501)	(−2.570)
4	−0.1071	−0.0663	−0.1115	0.0568	−0.1103
SU	(−2.840)	(−2.013)	(−2.883)	(2.140)	(−2.943)
MA	(−2.390)	(−1.665)	(−2.432)	(1.744)	(−2.527)
6	−0.0954	−0.0853	−0.1027	0.0988	−0.1052
SU	(−2.518)	(−2.346)	(−2.630)	(3.580)	(−2.803)
MA	(−2.117)	(−1.916)	(−2.227)	(2.881)	(−2.441)
8	−0.0779	−0.1060	−0.0891	0.1423	−0.0981
SU	(−2.034)	(−2.544)	(−2.252)	(4.837)	(−2.605)
MA	(−1.708)	(−2.052)	(−1.917)	(3.870)	(−2.301)
10	−0.0532	−0.1289	−0.0698	0.1878	−0.0884
SU	(−1.358)	(−2.626)	(−1.724)	(5.818)	(−2.330)
MA	(−1.138)	(−2.105)	(−1.480)	(4.672)	(−2.083)
12	−0.0188	−0.1549	−0.0428	0.2362	−0.0746
SU	(−0.455)	(−2.619)	(−1.014)	(6.435)	(−1.939)
MA	(−0.383)	(−2.103)	(−0.881)	(5.246)	(−1.753)
14	0.0300	−0.1855	−0.0044	0.2889	−0.0535
SU	(0.637)	(−2.548)	(−0.095)	(6.628)	(−1.352)
MA	(0.546)	(−2.067)	(−0.085)	(5.553)	(−1.235)

We also find that our results are essentially unaffected when $\tilde{\delta}_k$ and $\tilde{\Delta}_k^{BK}$ are jointly estimated with the VAR parameters and the covariance matrix of the residuals (together with its Choleski factor). For example, the estimate of $\tilde{\delta}_{CG}$ has a t-ratio of -8.947, while the estimate of Δ_{CG}^{BK} has a t-ratio of -2.179.

B. Some Further Tests

Table VI provides a further illustration of the how the tests based on our variance bounds perform relative to tests based on the H&J variance bound. Table VI reports estimates of $\tilde{\Delta}^{HJ}$, $\tilde{\delta}_k$, and $\tilde{\Delta}_k^{BK}$ (the positivity restriction is imposed) for the two risk variables y_{INF} and y_{CG}, with values of the relative-risk-aversion coefficient between 40 and 120. We consider the conditional consumption-based (normalized) pricing kernel with time separability, as in Section IV.A. The table clearly shows how our variance bounds are significantly more stringent than the H&J bound. Based on the H&J test, the lowest value of γ for which the consumption-based kernel cannot be rejected at the 5 percent level is between 40 and 60. On the other hand, as γ becomes larger, so does the shift in our variance bounds relative to the H&J bound. Based on our bounds, the lowest value of γ for which the consumption-based kernel cannot be rejected at the 5 percent level is between 60 and 80 for y_{INF}, and between 100 and 120 for y_{CG}.

A graphical illustration of the tests of Table VI is provided in Figures 1 and 2. Figure 1 plots the H&J variance bound (HJ) as a flat line and our variance bounds (BK) as a parabola. The risk variable is here y_{INF}, inflation risk. The figure also plots the locus of risk premium-variance pairs implied by the candidate (normalized) pricing kernel for the conditional time-separable case (C-CAPM) for different values of the relative-risk-aversion parameter γ. Figure 2 performs the same analysis for consumption risk. In both figures, we obtain the variance bounds on the pricing kernel by imposing the positivity restriction. $\tilde{\Delta}^{\mathrm{HJ}}$ is the vertical distance between the C-CAPM locus of risk premium-variance pairs and the H&J variance bound. $\tilde{\delta}_k$ is the vertical distance between our variance bounds and the H&J variance bound. $\tilde{\Delta}_k^{\mathrm{BK}}$ is the vertical distance between the C-CAPM locus and our variance bound.

These results increase the severity of Mehra and Prescott's (1985) equity-premium puzzle. As a possible solution of this puzzle, some articles (see, for example, Ferson and Constantinides (1991)) have introduced habit persistence in the candidate pricing kernels.

Table VII presents the results of tests of the unconditional (normalized) pricing kernel $x_{t+1} = \mathrm{IMRS}_{t+1}/\mathrm{E}(\mathrm{IMRS}_{t+1})$, for the case of habit persistence (Panel B) and, for comparison, for the time-separable case (Panel A). We report estimates of $\tilde{\Delta}^{\mathrm{HJ}}$, $\tilde{\delta}_k$, and $\tilde{\Delta}_k^{\mathrm{BK}}$ (positivity is imposed) for the two risk variables y_{INF} and y_{CG}, for values of the relative risk aversion coefficient between 30 and 110 for the time-separable case, and between 2 and 14 for the case of habit persistence. We set $\phi = -0.5$, which is the same value used in H&J and in the studies of the H&J regions by Gallant, Hansen, and Tauchen (1990), Cochrane and Hansen (1992), and Cecchetti, Lam, and Mark (1994). While this value of ϕ is lower than some of the estimates obtained by Ferson and Constantinides (1991), stronger habit persistence might lead to negative marginal utility. (This can be seen from the numerator and the denominator of equation (56).) For example, for our data set, when $\gamma = 12$ the marginal utility turns negative, in some instances, for values of ϕ lower than -0.61.

Also, we set $\rho = 1$ (the same value used in H&J, and in Gallant, Hansen, and Tauchen (1990) and Cochrane and Hansen (1992)[12]). Based on the H&J variance bound, the lowest value of γ for which the consumption-based model cannot be rejected at the 5 percent level is between 50 and 70 for the time-separable case, and between 8 and 10 for habit persistence. By comparison, in the case of the risk variable y_{CG}, the corresponding values for γ implied by our variance bounds are between 90 and 110 for the time-separable case, and between 12 and 14 for the habit persistence case. Moreover, the values of γ for which the candidate pricing kernel exactly meets the H&J restriction are 83 and 12.85 for the time-separable and habit-persistence cases, respectively. The values of γ for which the candidate pricing kernel meets our variance bounds are 95.10 and 14.18 for the risk variable y_{INF}, and 149.74 and 16.64 for the risk variable y_{CG}.

[12] Cecchetti, Lam, and Mark (1994) set ρ equal to 0.99 and 1.02.

As we expected, our results increase the severity of the equity-premium puzzle, and differ from those of existing statistical tests of the C-CAPM based on the H&J regions. Cochrane and Hansen (1992) find that a relative risk aversion of one (log utility) does not lead to a rejection of the time-separable consumption-based model at the 5 percent level, based on quarterly returns on stocks and bonds. And Cecchetti, Lam, and Mark (1994) find that a value of the relative-risk-aversion coefficient of one is at least marginally consistent (the model cannot be rejected at the 1 percent level) with monthly data on stock and bond returns for both the time-separable and habit-persistence cases.

This exacerbation of the equity-premium puzzle is the result of three elements. First, in our tests of the conditional (normalized) pricing kernel, we impose the conditional restriction $E_t(x_{t+1}) = 1$ (see the discussion in Section IV.A). Second, we consider the rate of return r_{HL} on the book-to-market factor, whose monthly Sharpe ratio (0.1540) is much higher than that of the market-portfolio proxy (0.0961) and implies especially stringent variance bounds. (Recall that the maximum Sharpe ratio obtainable from a set of assets determines a lower bound on the variability of a pricing kernel.) This result is consistent with Snow (1991) who also notes that the equity-premium puzzle becomes more severe when using a different choice of assets than those used by Hansen and Jagannathan (1991). Third, and this is the main implication of our article, our variance bounds are significantly more stringent than the H&J bound, especially for inflation and consumption risk.

VI. Conclusions

This article investigates the link between the variability of a pricing kernel and the risk premia that it assigns to arbitrary sources of risk, and follows the intuition of Hansen and Jagannathan (1991). We show that the lower bound on the variance of a pricing kernel increases with the (squared) difference between the premium that it assigns to a given risk variable and the premium on the mimicking portfolio. We find that the difference between our minimum-variance bounds and the H&J bound is the squared Sharpe ratio of a portfolio that is long the risk variable under consideration, and short its nearest hedge. Given the variance of a pricing kernel, we identify upper and lower bounds for the risk premium associated with a risk variable. This interval increases in size with the variance of the pricing kernel and with the variance of the risk variable that is explained by asset returns. It reduces to a single value when the risk variable can be exactly replicated by using available securities.

Our analysis has several appealing features that contrast with the existing literature on the subject. First, the validity of our minimum-variance and risk-premium bounds does not depend on the appropriate specification of an asset-pricing model and a return-generation process. Second, the risk variables that we consider do not need to be replicated by traded assets. Third, while we allow for time variation of the risk premia, our bounds do not require the explicit modeling of such time variation. Fourth, conditioning information can be easily incorporated to expand the set of assets under scrutiny, and sharpen the bounds.

As an application, our bounds on the risk premia and variance of a pricing kernel can be used as diagnostics for an explicit asset-pricing model. We consider the standard representative-agent consumption-based pricing kernels with constant relative risk aversion for the cases of time separability and habit persistence. We find that our variance bounds are significantly more stringent than the H&J bound (accounting for sampling error) for both inflation and consumption risk; and that considering consumption-growth risk leads to a marginal rejection of the consumption-based model, even for levels of relative risk aversion as high as 110 in the time-separable case, and as high as 12 in the habit-persistence case. This last result increases the severity of the equity-premium puzzle of Mehra and Prescott (1985). It also contrasts with the results of Cochrane and Hansen (1992) and Cecchetti, Lam, and Mark (1994), who find that a relative risk aversion coefficient of one is at least marginally consistent with data on stock and bond returns.

Appendix

A. Conditioning Information

Consider again the conditional restrictions:

$$\mathrm{E}_t(q_{t+1}r_{t+1}) = \mathbf{1}r_{ft} \tag{A1}$$

$$\mathrm{E}_t(q_{t+1}) = 1. \tag{A2}$$

Consider now a vector of instruments z_t, whose realization belongs to the information set at time t. Taking Kroneker products, we obtain:

$$\mathrm{E}_t[(q_{t+1}r_{t+1}) \otimes z_t] = \mathbf{1}r_{ft} \otimes z_t \tag{A3}$$

$$\mathrm{E}_t(q_{t+1} \otimes z_t) = \mathbf{1} \otimes z_t. \tag{A4}$$

Let $r^z_{a,t+1} \equiv [r^\top_{a,t+1} \otimes z^\top_t, z^\top_t]^\top$ and $\iota^z_t \equiv [(\mathbf{1} \otimes z_t)^\top r^\top_{ft}, z^\top_t]^\top$. Assuming stationarity and applying the law of iterated expectations to the moment conditions (A3) and (A4), we obtain

$$\mathrm{E}(qr^z_a) = \mathrm{E}(\iota^z). \tag{A5}$$

Hence, the bounds on the variance of the pricing kernel and on the risk premia obtained in (24) and (27) still hold, but the coefficient vectors α and α_{y_k} are replaced by:

$$\alpha^z = [\mathrm{E}(r^z_a r^{z\top}_a)]^{-1}\mathrm{E}(\iota^z) \tag{A6}$$

$$\alpha^z_{y_k} = [\mathrm{E}(r^z_a r^{z\top}_a)]^{-1}\mathrm{E}(y_k r^z_a). \tag{A7}$$

B. Positivity of the Pricing Kernel

Considering the mean risk premium λ_k, we can write:

$$\lambda_k \equiv -\mathrm{E}(qy_k) \tag{A8}$$

$$= -\mathrm{E}[(\tilde{q} + (q - \tilde{q}))(y_k^* + (y_k - y_k^*))] \tag{A9}$$

$$= -\mathrm{E}(\tilde{q}y_k^*) - \mathrm{E}[\tilde{q}(y_k - y_k^*)] - \mathrm{E}[(q - \tilde{q})y_k^*] - \mathrm{E}[(q - \tilde{q})(y_k - y_k^*)]. \tag{A10}$$

Since $\mathrm{E}(qr_a) = \mathrm{E}(\tilde{q}r_a)$, we have $\mathrm{E}[(q - \tilde{q})r_a] = 0$ and $(q - \tilde{q})$ is orthogonal to r_a. Hence, $\mathrm{E}[(q - \tilde{q})y_k^*] = \mathrm{E}[(q - \tilde{q})r_a\alpha_{y_k}^\top] = 0$, but, in general, $\mathrm{E}[\tilde{q}(y_k - y_k^*)] \neq 0$. Moreover, we have $\tilde{\lambda}_k \equiv -\mathrm{E}(\tilde{q}y_k) = -\mathrm{E}(\tilde{q}y_k^*) - \mathrm{E}[\tilde{q}(y_k - y_k^*)] = \lambda_k^* - \mathrm{E}[\tilde{q}(y_k - y_k^*)]$. Hence, we can write (A10) as $\lambda_k = \tilde{\lambda}_k - \mathrm{E}[(q - q^*)(y_k - y_k^*)]$. Rearranging, we obtain

$$\mathrm{E}[(q - \tilde{q})(y_k - y_k^*)] = \tilde{\lambda}_k - \lambda_k. \tag{A11}$$

By the Cauchy-Schwarz inequality, we have

$$|\mathrm{E}[(q - \tilde{q})(y_k - y_k^*)]| \leq \sqrt{\mathrm{E}[(q - \tilde{q})^2]\mathrm{E}[(y_k - y_k^*)^2]}. \tag{A12}$$

Moreover, we have:

$$\mathrm{E}[(q - \tilde{q})^2] = \mathrm{E}(q^2) + \mathrm{E}(\tilde{q}^2) - 2\mathrm{E}(q\tilde{q}) \tag{A13}$$

$$= \mathrm{E}(q^2) - \mathrm{E}(\tilde{q}^2) - 2\mathrm{E}[(q - \tilde{q})\tilde{q}] \tag{A14}$$

$$= \mathrm{Var}(q) - \mathrm{Var}(\tilde{q}) - 2\mathrm{E}[(q - \tilde{q})\tilde{q}]. \tag{A15}$$

Squaring both sides of equation (A12), using (A11), and rearranging, we obtain:

$$\mathrm{Var}(q) \geq \mathrm{Var}(\tilde{q}) + 2\mathrm{E}[(q - \tilde{q})\tilde{q}] + \frac{[\mathrm{E}((q - \tilde{q})(y_k - y_k^*))]^2}{\mathrm{E}[(y_k - y_k^*)^2]} \tag{A16}$$

$$= \mathrm{Var}(\tilde{q}) + 2\mathrm{E}[(q - \tilde{q})\tilde{q}] + \frac{(\tilde{\lambda}_k - \lambda_k)^2}{\mathrm{E}[(y_k - y_k^*)^2]}. \tag{A17}$$

and since $\mathrm{E}[(q - \tilde{q})\tilde{q}] = \mathrm{E}(q\tilde{q}) - \mathrm{E}(\tilde{q}^2) \geq 0$ (see equation (31)), this implies

$$\mathrm{Var}(q) \geq \mathrm{Var}(\tilde{q}) + \frac{(\tilde{\lambda}_k - \lambda_k)^2}{\mathrm{E}[(y_k - y_k^*)^2]} \tag{A18}$$

Consider now the case of several risk variables. Since $q = (q - \tilde{q}) + \tilde{q}$, we have $\mathrm{Var}(q) = \mathrm{Var}(q - \tilde{q}) + \mathrm{Var}(\tilde{q}) + 2\mathrm{Cov}(q - \tilde{q}, \tilde{q})$. Since $\mathrm{E}(qy^*) = \mathrm{E}(\tilde{q}y^*)$, we also have $-\mathrm{E}[(q - \tilde{q})y] = -E[(q - \tilde{q})(y - y^*)] = \tilde{\lambda}$, where $\tilde{\lambda}$ is the vector of $\tilde{\lambda}_k$. Thus, we obtain

$$\begin{aligned} \mathrm{Var}(q - \tilde{q}) &\geq (\lambda - \tilde{\lambda})^{\top}[\mathrm{E}((y - y^*)(y - y^*)^{\top})]^{-1}(\lambda - \tilde{\lambda}) \\ &= \sum_{k=1}^{K} \frac{(\tilde{\lambda}_k - \lambda_k)^2}{\mathrm{E}[(y_k - y_k^*)^2]}. \end{aligned} \tag{A19}$$

Using the results above, and the fact that $\mathrm{Cov}(q - \tilde{q}, \tilde{q}) = \mathrm{E}[(q - \tilde{q})\tilde{q}]$, we obtain the inequality

$$\mathrm{Var}(q) \geq \mathrm{Var}(\tilde{q}) + \sum_{k=1}^{K} \frac{(\tilde{\lambda}_k - \lambda_k)^2}{\mathrm{E}[(y_k - y_k^*)^2]} \tag{A20}$$

which is the analog of equation (26).

The analysis in the article focuses on the inequality in equation (A18), rather than on (A17). We do this for two reasons: First, the term $\mathrm{E}[(q - \tilde{q})\tilde{q}]$ in (A17) is not directly related to the variance of a pricing kernel nor to the risk premia assigned by a pricing kernel, which are the main objects of our analysis. Second, when we consider an explicit asset-pricing model, we replace q with a candidate (normalized) kernel x, which may not satisfy the inequality $\mathrm{E}(x\tilde{q}) \geq \mathrm{E}(\tilde{q}^2)$. Thus, the term $\mathrm{E}[(x - \tilde{q})\tilde{q}]$ may be negative (and substantial), and the empirical counterpart of the inequality (A17) may turn out to be less restrictive than the H&J bound.

Rearranging equation (A18), and taking the square root of both sides, we obtain

$$\begin{aligned} &\tilde{\lambda}_k - \sqrt{[\mathrm{Var}(q) - \mathrm{Var}(\tilde{q})]\mathrm{E}[(y_k - y_k^*)^2]} \\ &\qquad \leq \lambda_k \leq \tilde{\lambda}_k + \sqrt{[\mathrm{Var}(q) - \mathrm{Var}(\tilde{q})]\mathrm{E}[(y_k - y_k^*)^2]}. \end{aligned} \tag{A21}$$

REFERENCES

Abel, Andrew B., 1990, Asset prices under habit formation and catching up with the Joneses, *American Economic Review* 80, 38–42.

Burmeister, Edwin, and Marjorie B. McElroy, 1988, Joint estimation of factor sensitivities and risk premia for the arbitrage pricing theory, *Journal of Finance* 43, 721–733.

Burnside, Craig, 1994, Hansen-Jagannathan bounds as classical tests of asset-pricing models, *Journal of Business and Economic Statistics* 12, 57–79.

Cecchetti, Stephen G., Pok-Sang Lam, and Nelson C. Mark, 1994, Testing volatility restrictions on intertemporal marginal rates of substitution implied by Euler equations and asset returns, *Journal of Finance* 49, 123–152.

Chen, Zhiwu, and Peter J. Knez, 1994, A pricing operator-based testing foundation for a class of factor pricing models, *Mathematical Finance* 4, 121–141.

Chen, Nai-Fu, Richard Roll, and Stephen A. Ross, 1986, Economic forces and the stock market, *Journal of Business* 59, 383–403.

Cochrane, John H., 1992, Explaining the variance of price-dividend ratios, *Review of Financial Studies* 5, 243–280.

Cochrane, John H., and Lars Peter Hansen, 1992, Asset pricing explorations for macroeconomics, in Olivier J. Blanchard and Stanley Fischer, *NBER Macroeconomic Annual 1992*, 7, 115–165.

Cochrane, John H., and Jesús Saá-Requejo, 1996, Beyond arbitrage: Good-deal pricing of derivatives in incomplete markets, Working paper, University of Chicago.

De Santis, Giorgio, 1995, Volatility bounds for stochastic discount factors: Tests and implications from international stock returns, Working paper, University of Southern California.

Downs, David H., and Karl N. Snow, 1994, Sufficient conditioning information in dynamic asset pricing, Working paper, University of North Carolina.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Ferson, Wayne E., and George M. Constantinides, 1991, Habit persistence and durability in aggregate consumption: Empirical tests, *Journal of Financial Economics* 29, 199–240.

Ferson, Wayne E., and Campbell R. Harvey, 1991, The variation of economic risk premia, *Journal of Political Economy* 99, 385–415.

Gallant, A. Ronald, Lars Peter Hansen, and George Tauchen, 1990, Using conditional moments of asset payoffs to infer the variance of intertemporal marginal rates of substitution, *Journal of Econometrics* 45, 141–179.

Grossman, Sanford J., and Robert J. Shiller, 1981, The determinants of the variability of stock market prices, *American Economic Review* 71, 222–227.

Hansen, Lars Peter, 1982, Large sample properties of generalized method of moments estimators, *Econometrica* 50, 1029–1054.

Hansen, Lars Peter, John Heaton, and Erzo Luttmer, 1995, Econometric evaluation of asset pricing models, *Review of Financial Studies* 8, 237–274.

Hansen, Lars Peter, and Ravi Jagannathan, 1991, Implications of security market data for models of dynamic economies, *Journal of Political Economy* 99, 225–262.

Hansen, Lars Peter, and Ravi Jagannathan, 1997, Assessing specification errors in stochastic discount factors models, *Journal of Finance* 52, 557–590.

Hansen, Lars Peter, and Scott F. Richard, 1987, The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models, *Econometrica* 55, 587–613.

Hansen, Lars Peter, and Kenneth J. Singleton, 1982, Generalized instrumental variables estimation of nonlinear rational expectations models, *Econometrica* 50, 1269–1286.

Harrison, J. Michael, and David M. Kreps, 1979, Martingales and arbitrage in multiperiod securities markets, *Journal of Economic Theory* 20, 381–408.

Knez, Peter J., 1994, Pricing money market securities with stochastic discount factors, Working paper, University of Wisconsin-Madison.

MacKinlay, A. Craig, 1995, Multivariable models do not explain deviations from the CAPM, *Journal of Financial Economics* 38, 3–28.

Mehra, Rajnish, and Edward C. Prescott, 1985, The equity premium: A puzzle, *Journal of Monetary Economics* 15, 145–161.

Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

Snow, Karl N., 1991, Diagnosing asset pricing models using the distribution of asset returns, *Journal of Finance* 46, 955–983.