Linear Approximations and Tests of Conditional Pricing Models *

Michael W. Brandt†
Duke University and NBER

David A. Chapman‡
Boston College

May 2014

Abstract

If a nonlinear risk premium in a conditional asset pricing model is approximated with a linear function, as is commonly done in empirical research, the fitted model is misspecified. We use a generic reduced-form model economy with modest risk premium nonlinearity to examine the size of the resulting misspecification-induced pricing errors. Pricing errors from modest nonlinearity can be large, and a version of a test for nonlinearity based on risk premiums rather than pricing errors has reasonable power properties after properly controlling for the size of the test. We conclude by examining the importance of modest nonlinearity in the recent structural general equilibrium model of Papanikolaou (2011).

*We thank Keith Brown, Jennifer Conrad, Joost Driessen, Jin-Chuan Duan, Wayne Ferson, Xavier Gabaix, Lorenzo Garlappi, Eric Ghysels, Larry Harris, Kevin Huang, Tim Johnson, Frank de Jong, Mark Kamstra, Raymond Kan, Ed Kane, Jon Lewellen, Andy Lo, Ludovic Phalippou, Jacob Sagi, Robert Stambaugh, and seminar participants at various universities and symposia.

†Fuqua School of Business, 100 Fuqua Drive, Box 90120, Durham, NC 27708. email: `mbrandt@duke.edu`.

‡Carroll School of Management, Fulton 330, Boston College, 140 Commonwealth Avenue, Chestnut Hill, MA 02467, e-mail: `david.chapman@bc.edu`.

1 Introduction

The empirical evidence against unconditional versions of the classic capital asset pricing model (CAPM) and the consumption CAPM (CCAPM), gathered over the past three decades, has inspired a large literature that examines conditional versions of these models; examples of this line of research include Jagannathan and Wang (1996), Lettau and Ludvigson (2001), Lustig and van Nieuwerburgh (2005), Santos and Veronesi (2006), and Yogo (2006). In many conditional asset pricing models, loadings of asset returns on common risk factors and the associated factor risk premiums vary over time as some function of observable state variables. We examine how inference about conditional asset pricing models is affected by misspecification of the functional form of time-varying risk premiums.

Conditional models can differ from their unconditional counterparts through time varying factor loadings and/or time-varying risk premiums. While the relative importance of these alternate channels depends on the choice of factors and on the functional form of conditioning in the model, Ferson and Harvey (1991) argue that, in a standard linear macro-factor model, time-variation in risk premiums is more important than time-variation in the factor loadings for explaining observed predictability in excess returns. Lewellen and Nagel (2006) conclude, using a methodology based on high-frequency estimation, that time-varying loadings cannot be a source of the empirical success of a conditional form of the CAPM relative to the unconditional CAPM in a cross-section of expected returns.

In a conditional asset pricing model, the form of any time-variation depends on the joint evolution of the underlying data generating process and the economic agents' information sets. These aspects of the theoretical model are unspecified in a reduced-form theory, and they are generally unobservable; see the discussion in Hansen and Richard (1987). The econometrician must, therefore, model the time-variation in a way that is tractable. In the case of time-varying factor loadings, covariances are generally specified as functions of past returns and factor innovations, following the ARCH/GARCH modeling framework, or more simply as linear or exponentially linear functions of state-variables that are observable by the econometrician. In the case of time-varying risk premiums, the vast majority, if not all applications involve linear specifications in observable state-variables. This restriction seems, a priori, less credible given newer structural models that allow for technological innovation and financial sector feedback to the real economy that are inherently nonlinear.[1]

How sensitive are standard econometric tests of conditional asset pricing models to the assumption that time-variation in loadings and risk premiums are linear functions of observable state variables? Of course, this question can only be answered within the context of a specific assumption about the form of the true underlying risk premiums. The question is partially addressed by Ghysels (1998), who shows that when the factor loadings of a conditional pricing model exhibit structural breaks, testing the model with linear specifications can lead to statistically less reliable inference than testing an unconditional version of the model. This is true despite the fact that the data is generated by a conditional pricing model. It is reasonable to consider time varying loadings first, since variances and covariances are estimated with much greater precision than expected returns. There has, however, been no comparable analysis examining the sensitivity of statistical inferences to the linearity assumption for conditional risk premiums. Our paper fills this gap in the literature.

We show that, under the assumption that risk premiums can be well-approximated by a polynomial function of the underlying state variables, misspecification of the risk premiums can be reformulated as a missing-variables problem. This reformulation yields analytical expressions for the misspecification-induced mispricing as well as test statistics that can be used to detect misspecification. We then examine a reduced-form conditional asset pricing model that features a modest and realistic degree of nonlinearity in the risk premiums. The model is an extension of the intertemporal CAPM specified in Brennan, Wang, and Xia (2004). The pricing kernel of this model depends on innovations to the market portfolio, innovations to the short rate, and innovations to the Sharpe ratio of the tangency portfolio. It is a general representation for a large class of models that imply a time-varying capital market line.

We calibrate the model to size and book-to-market sorted portfolio returns, a short-term Treasury yield, and estimates of the Sharpe ratio of the S&P 500 index, and we examine the analytical and simulated pricing errors caused by misspecifying the risk premiums as linear functions of the state variables as well as the finite-sample properties of alternative tests for nonlinear risk premiums. Our simulations show that misspecification-induced pricing errors can

[1] As an additional example, Roussanov (2012) uses nonparametric estimation techniques to demonstrate the empirical significance of nonlinear risk premiums in a conditional consumption-based model that includes status effects.

be large both statistically and relative to average excess returns. Given our use of the stochastic discount factor (SDF) representation, there are standard tests for the importance of omitted nonlinear terms. One test is based directly on the magnitude of the pricing errors. The other test examines the estimated risk premiums on the higher-order terms directly. We demonstrate that it is important to correct the size of the nonlinearity tests in realistic samples. In our tests, the nonlinearity test based on higher-order risk premium terms rather than on pricing errors has better finite sample power to detect modest nonlinearity.

As a final step in our analysis, we examine modest nonlinearity in the model of embodied technological change introduced in Papanikolaou (2011). When the model is estimated over a long sample of quarterly data on the standard 25 size- and book-to-market-sorted portfolios from 1948 to 2012 using the standard linear approximation, there is limited evidence that the model can explain the cross-section of expected returns. However, adding the nonlinear conditional premiums terms that are a direct implication of the model results in a significant improvement in the cross-sectional explanatory power of the model, and it provides substantial evidence in favor of the importance of embodied technological change in explaining measured return differences.

2 Conditional Asset Pricing Models

2.1 Pricing Models in the Stochastic Discount Factor Form

Standard practice for estimating and testing conditional pricing models has evolved over time in a series of papers that include Campbell (1987), Gibbons and Ferson (1985), Harvey (1989), Shanken (1990), and Cochrane (1996), culminating in the textbook treatment of Cochrane (2005). In order to confront the standard testing methodology with a generic pricing model, we consider a class of economies with complete financial markets and no arbitrage opportunities. The general form of the pricing kernel, under the physical measure, is:

$$M_{t+1} = \exp\left(r_t^f - \frac{1}{2}\boldsymbol{\Lambda}_t'\boldsymbol{\Lambda}_t + \boldsymbol{\Lambda}_t'\varepsilon_{t+1}\right), \tag{1}$$

where r_t^f is the one-period risk-free return, $\boldsymbol{\Lambda}_t$ is a K-vector of pricing kernel functions, and $\varepsilon_{t+1} \sim \mathcal{N}\left(\mathbf{0}, \mathbf{I}_K\right)$ is a vector of normalized common factor innovations. The time varying nature of $\mathbf{\Lambda}_t$ makes this a conditional pricing model.

The exponential-affine structure of (1) implies that the log pricing kernel is linear in the factors. We can rewrite (1) as

$$m_{t+1} \equiv \ln M_{t+1} = r_t^f - \frac{1}{2}\mathbf{\Lambda}_t'\mathbf{\Lambda}_t + \mathbf{\Lambda}_t'\varepsilon_{t+1}. \tag{2}$$

Linearity in the normalized factor innovations has a long history in financial economics, as either a direct assumption about the pricing kernel, see Chen, Roll, and Ross (1986), as an implication of the capital asset pricing model of Sharpe, Lintner, and Mossin, or as a first-order approximation to a general consumption-based pricing model; see Breeden, Gibbons, and Litzenberger (1989) or Lettau and Ludvigson (2001). The form of (1) is also common in the literature on dynamic term structure models; e.g., Dai, Singleton, and Yang (2007).

The multifactor pricing kernel in (2) is equivalent – in a pricing sense – to a kernel that is linear in the return to the (conditionally) mean variance efficient portfolio formed by projecting (2) onto the space of marketed asset payoffs; see Hansen and Richard (1987). This (conditionally) efficient portfolio defines a unique (if the kernel is unique) maximum Sharpe ratio process, denoted S_t. In the case of the kernel in (1), this implies

$$S_t = \frac{\sigma_t\left(M_{t+1}\right)}{E_t\left(M_{t+1}\right)} = \sqrt{\exp\left(\mathbf{\Lambda}_t'\mathbf{\Lambda}_t\right) - 1}, \tag{3}$$

where the first equality follows from the definition of the Sharpe ratio and the second equality follows from the (conditional) log-normality of the pricing kernel.[2]

At this level of generality, the pricing kernel innovations have no specific economic interpretation. They could represent innovations to aggregate consumption or a function of aggregate consumption, as in a habit-formation model, innovations to an aggregate market portfolio, or both. The values in the vector ε_{t+1} can even be innovations to the risk premium functions themselves. The normality assumption for the shocks is an important simplification, but the unconditional distribution of the pricing kernel can still be fat-tailed and/or skewed depending on the unconditional distribution of $\mathbf{\Lambda}_t$. The pricing kernel coefficients, $\mathbf{\Lambda}_t$, can be interpreted as risk premiums

[2] The proof of (3) is straightforward. The first equality simply manipulates the definition of a stochastic discount factor and a conditional correlation coefficient. The second equality follows immediately from the conditional log-normality of the pricing kernel in (2).

because ε_{t+1} are normalized to be uncorrelated and have unit variance. We will refer to $\mathbf{\Lambda}_t$ as risk premium functions throughout the remainder of the paper.

The standard assumption used in the empirical literature is that the elements of Λ_t are linear functions of information available to the market at time t.[3] This may be a reasonable assumption, but it is not easily justified for different equilibrium models, since these models are frequently silent on the precise functional form of conditional risk premiums. Indeed, Bansal, Hsieh, and Viswanathan (1993) and Chapman (1997) provide examples of both returns and consumption-based models with nonlinearities in the risk vs. return trade-off. It seems reasonable that significant nonlinearities in Λ_t would be important – but easily detected – specification errors. However, what about modest levels of nonlinearity? Can moderate nonlinearity be detected in noisy realized returns? Does moderate nonlinearity have an economically significant effect on measured pricing errors?

2.2 Alphas Generated by Misspecification

In evaluating the effect of misspecification, we assume that there is a single risk premium and that it can be well-approximated by a set of orthonormal polynomials[4]

$$\Lambda\left(Z_t\right) = \psi_0 + \psi_1 Z_t + P_2\left(Z_t\right) + P_3\left(Z_t\right) + \cdots , \tag{4}$$

where $P_j\left(Z_t\right)$ is the component of Z_t^j that is orthogonal to Z_t^i and $P_i\left(Z_t\right)$, for $i < j$, and normalized to have zero mean and unit variance.[5] We further assume, consistent with standard practice, a linear structure for the asset pricing kernel:

$$M_{t+1} = 1 + \mathbf{f}_{t+1}'\mathbf{b}_t. \tag{5}$$

[3] An early example of this assumption is Shanken (1990); see also Cochrane (2005) for a textbook discussion.

[4] The assumption of one versus multiple risk premiums is made purely for notational convenience. Our argument extends naturally to multiple risk premiums. The polynomial expansion assumption is critical for our omitted variables interpretation. However, there is a large literature in numerical analysis on the use of polynomial approximations to general functions; see Chapter 6 in Judd (1998).

[5] See Chapman (1997) for another example of the use of orthonormal polynomial approximations in an asset pricing setting.

The information used in determining asset prices is reduced to a K-vector of observable variables, $\mathbf{Z}_t$, and we write the pricing kernel coefficient, $\mathbf{b}_t$, as

$$\mathbf{b}_t = \mathcal{B}\mathbf{Z}_t, \tag{6}$$

where $\mathcal{B}$ is a $K \times L$ matrix of constant coefficients.

Under this approximation, the linearity imposed by equation (6) implies the omission of the higher order terms $P_j(Z_t)$ for $j \geq 2$. Given the simple structure of the pricing kernel implied by standard estimation methods for linear models, it is then possible to compute a closed-form expression for the mispricing, or alphas, generated by misspecification.[6] These alphas are based on the true moments of the factors and excess returns. Let

$$\mathbf{F}_{t+1} = \left(\mathbf{F}'_{1,t+1}, \mathbf{F}'_{2,t+1}\right)', \tag{7}$$

where $\mathbf{F}_1$ is the L_1-vector of linear terms in the conventional approximation and $\mathbf{F}_2$ is the L_2-vector of higher-order terms omitted by the approximation, where $L = L_1 + L_2$.[7] The covariance matrix of the factors can be partitioned accordingly.[8] It is important to note that – by construction of the higher order approximation – the omitted variables are orthogonal to the included factors, which significantly simplifies the analysis.

In the true model, expected excess returns are exactly linear in the betas with respect to $\mathbf{F}$, and true pricing errors are zero. If the market prices of risk are partitioned conformably, then

$$E\left(\mathbf{R}^{e\prime}\right) = \left[\begin{array}{ccc} \lambda'_1 & \vdots & \lambda'_2 \end{array}\right]\mathbf{B}, \tag{8}$$

where

$$\mathbf{B} = \mathbf{\Sigma}_{\mathbf{F}}^{-1}\mathbf{\Sigma}_{\mathbf{FR}^e}, \tag{9}$$

and $\mathbf{\Sigma}_{\mathbf{FR}^e}$ is the $L \times N$ covariance matrix of the $\mathbf{F}$ factors with the N returns in $\mathbf{R}^e$. In equation (8), expected returns are expressed exactly as a function of the factors that are included in the

[6]See Appendix A for details.

[7]Even though there is a single (composite) factor earning a risk premium in the true conditional model, there are multiple factors in the unconditional model used as an approximation.

[8]See Appendix B for details.

conventional approximation, the factors that are excluded from the conventional approximation, and the covariance of these two components. Since $\Sigma_{12} = \Sigma_{21} = 0$, by construction, we can rewrite (8) as

$$E\left(\mathbf{R}^{e}\right) = \mathbf{\Sigma}_{\mathbf{F}_1\mathbf{R}^e}\mathbf{\Sigma}_{11}^{-1}\lambda_1 + \mathbf{\Sigma}_{\mathbf{F}_2\mathbf{R}^e}\mathbf{\Sigma}_{22}^{-1}\lambda_2, \tag{10}$$

where $\mathbf{\Sigma}_{11}^{-1}$is the inverse of the partition of the covariance matrix associated with the included factors and $\mathbf{\Sigma}_{22}^{-1}$ is the inverse of the partition of the covariance matrix associated with the excluded factors. It follows immediately that the pricing errors, alphas, induced by ignoring nonlinearity in the risk premium function are:

$$\alpha = \mathbf{\Sigma}_{\mathbf{F}_2\mathbf{R}^e}\mathbf{\Sigma}_{22}^{-1}\lambda_2. \tag{11}$$

In order for these pricing errors to be economically significant, the higher order terms must covary strongly with excess returns and their risk prices must be large.

2.3 Testing for Nonlinearity

The first test for the statistical significance of nonlinearity in risk premiums is based on the estimates of alphas constructed from the unrestricted model (the specification that includes second-order terms) and the restricted (linear) model:

$$T\left(\widehat{\alpha}_R - \widehat{\alpha}_U\right)' \widehat{\mathbf{\Sigma}}_{\alpha_R}^{-1}\left(\widehat{\alpha}_R - \widehat{\alpha}_U\right) \overset{a}{\sim} \chi_q^2, \tag{12}$$

where $\widehat{\alpha}_U$ is the N-vector of alpha estimates under the unrestricted model, $\widehat{\alpha}_R$ is the N-vector of alphas under the restricted model, and $\widehat{\mathbf{\Sigma}}_{\alpha_R}$ is the estimated covariance matrix of alphas under the restricted model (see Appendix B). This statistic is analogous to the GMM-type test statistic based on differences in the value of the objective function under the two models, and we will refer to it below as the "α-test."

If the higher-order risk premium terms are important for pricing the test assets, then their inclusion should result in a large reduction in estimated alphas, and the test statistic in (12) should be large. It is important to compare the distance between the alpha vectors using a common covariance matrix. This ensures that the test statistic is non-negative.

The second test statistic is a simple quadratic form in the estimated risk premiums:

$$\widehat{\lambda}_2' \widehat{\mathbf{\Sigma}}_{\lambda_2}^{-1} \widehat{\lambda}_2 \overset{a}{\sim} \chi_q^2, \tag{13}$$

where $\widehat{\lambda}_2$ is the q-vector of estimates of higher-order risk premiums and $\widehat{\mathbf{\Sigma}}_{\lambda_2}$ is the estimated covariance matrix of these risk premiums. Estimates of the higher-order risk premium terms that are large relative to their estimated variances and covariances are inconsistent with the null hypothesis that a linear approximation is sufficient. We will refer to this test as the "λ-test."

The α-test and λ-test follow the common practice in the empirical literature of assuming that the estimated model is correctly specified. Kan, Robotti, and Shanken (2012) (hereafter KRS) construct a test for nested models based on the cross-sectional R^2 statistic from the second-pass regression that is robust to misspecification.[9] The (nested version of the) "*KRS*-test" of $H_0 : \hat{\rho}_U^2 = \hat{\rho}_R^2$ uses the fitted (cross-sectional) R^2 from each model, $\hat{\rho}_i^2$ for $i \in \{U, R\}$. It is defined as:

$$T\left(\hat{\rho}_U^2 - \hat{\rho}_R^2\right) \overset{a}{\sim} \sum_{j=1}^{L_2} \frac{\iota_j}{Q_0} \tilde{x}_j, \tag{14}$$

where Q_0 is a quadratic form in the weighting matrix (e.g., OLS or GLS) and the cross-sectional deviations in expected returns, ι_j is the j-th eigenvalue of the matrix product of the inverse of the appropriate partition of the model parameter covariance matrix and the covariance matrix of the additional risk premium terms in the unrestricted model, and $\tilde{x}_j$ are independent χ_1^2 random variables.[10]

This distribution is dependent on the model (through the estimated weights $\frac{\iota_j}{Q_0}$) and the weighting matrix (OLS vs. GLS). However, it can be computed via simulation. As KRS note, the model ranking based on the differences in R^2 just re-arranges the information in the quadratic form based on the risk premiums, but there is potential value in examining the finite-sample properties of nonlinearity tests based on misspecification-robust versions of this statistic.

[9] KRS also construct the more general test statistic for comparing non-nested models, but that is not relevant to our analysis.

[10] The asymptotic distribution of the statistic in (14) is derived in Propositions A.5 and A.6 in the internet appendix to KRS.

3 A Calibrated Example

3.1 The Pricing Kernel

This example follows the general approach of Brennan, Wang, and Xia (2004) (hereafter, BWX) for implementing a version of the Intertemporal Capital Asset Pricing Model (ICAPM) of Merton (1973). We differ from BWX by using a discrete-time formulation of the economy and, more importantly, in the specification, identification, and calibration of the dynamics of the state variables.

Let r_t^f denote the yield on a one-period, default-risk free bond from t to $t+1$. The dynamics of this short rate are exogenously specified as

$$\ln r_t^f = \alpha_1 + \beta_1 \ln r_{t-1}^f + \varepsilon_{1t}, \tag{15}$$

where $\varepsilon_{1t} \overset{iid}{\sim} \mathcal{N}\left(0, \sigma_1^2\right)$. (15) was introduced into the term structure literature in Black, Derman, and Toy (1990) and Black and Karasinski (1991). The dynamics of the log of the maximum Sharpe ratio process are given by

$$\ln S_t = \alpha_2 + \beta_2 \ln S_{t-1} + \gamma \ln r_t^f + \varepsilon_{2t}, \tag{16}$$

where $\varepsilon_{2t} \overset{iid}{\sim} \mathcal{N}\left(0, \sigma_2^2\right)$. By construction, (15) and (16) force the conditional correlation between ε_{1t} and ε_{2t} to equal zero, for all t.

Equations (15) and (16) are a linear, first-order vector autoregression in logs. (16) implies that the Sharpe ratio on the conditionally mean-variance efficient portfolio can never be negative. There is no specific structural model underlying (16), but there are well-known business cycle related movements in both short-term interest rates and estimated Sharpe ratios that are consistent with a nonzero γ. We will estimate the parameters of (15) and (16) below.

Let R_{t+1}^M denote the return on a broad market portfolio, and let $\varepsilon_{3t} \overset{iid}{\sim} \mathcal{N}\left(0, \sigma_3^2\right)$ denote the innovation, at t, to this return. The unique pricing kernel for the economy, from t to $t+1$, under the physical measure, is defined as a special case of the kernel in equation (1):

$$M_{t+1} = \exp\left(r_t^f - \frac{1}{2}\Lambda_t^2 + \Lambda_t \omega^{-1} \xi_{t+1}\right), \tag{17}$$

where Λ_t is a risk premium parameter, $\xi_{t+1} \equiv \delta' \varepsilon_{t+1}$, $\varepsilon_{t+1} = (\varepsilon_{1t+1}, \varepsilon_{2t+1}, \varepsilon_{3t+1})'$, δ is a vector of constants, $\omega = (\delta' \Sigma \delta)^{1/2}$, and $\Sigma = I_3$ is the covariance matrix of ε_{t+1}. The single innovation to the kernel is a weighted-average of the innovations to the market return, the risk-free rate, and the maximum Sharpe ratio.[11] This is the discrete-time analog of the specification in BWX. The fact that the shock, ξ, is normally distributed implies that M_t is conditionally lognormal.

There are three economically significant restrictions imposed on the pricing kernel by (17). First, we assume a single risk premium on a single composite shock. This is a simplifying assumption. The basic issue of the efficacy of the standard linear conditioning approximation carries over into a multifactor setting. Second, the market portfolio plays a specific role in the pricing kernel. This is true in a variety of common models (the CAPM, Fama and French, 1993, and Epstein and Zin, 1989, are examples), but it is not true of all prominent pricing models. Campbell and Cochrane (1999), Bansal and Yaron (2004), and Lustig and van Nieuwerburgh (2005) are all examples of recent consumption-based models in which the market portfolio (and its innovations) plays no distinct role in the pricing kernel. Third, the Sharpe ratio is a nonlinear function of the risk premium, as noted in equation (3), which means that the pricing kernel in (17) assumes that innovations to the risk premiums help determine the pricing kernel directly.

The final assumption we make in specifying the example economy is that from t to $t+1$ the asset returns and pricing kernel have a joint lognormal distribution, conditional on the current realizations of the market and the factors. Under this assumption, the fundamental asset pricing equation implies a generalization of the basic moment conditions in Hansen and Singleton (1983), where marginal utility of consumption growth, in that setting, is replaced by the pricing kernel:

$$E_t\left(r_{i,t+1}\right) + E_t\left(m_{t+1}\right) + \frac{1}{2}\left[\mathrm{var}_t\left(r_{i,t+1}\right) + \mathrm{var}_t\left(m_{t+1}\right) + 2\mathrm{cov}_t\left(r_{i,t+1}, m_{t+1}\right)\right] = 0, \tag{18}$$

for $i = 1, \ldots, N$, where $r_{i,t+1} \equiv \ln R_{i,t+1}$ and $R_{i,t+1}$ is the gross return to asset i. Equation (18) can be rewritten in terms of returns, functions of latent factors, and factor innovations:

$$E_t\left(r_{i,t+1}\right) + \frac{1}{2}\mathrm{var}_t\left(r_{i,t+1}\right) = r_t^f + \Lambda_t \mathrm{cov}\left(r_{i,t+1}, \omega^{-1}\xi_{t+1}\right), \tag{19}$$

[11] In Section 3.2, below, we show that δ cannot be identified uniquely from the time series regression of excess returns on the (extended) set of factors. However, the values of δ that are implicitly chosen in the regression in (25) are those that best fit the set of test assets used below.

for $i = 1, \ldots, N$. The covariance in equation (19) is unconditional given the assumptions on the state variable innovations that serve as factors.

Given that the factor, $\omega^{-1}\xi_{t+1}$, is *iid* standard normal, equation (19) can be rewritten in a traditional single-factor (conditional) beta pricing form:

$$E_t\left(r_{i,t+1}\right) + \frac{1}{2}\mathrm{var}_t\left(r_{i,t+1}\right) - r_t^f = \Lambda_t \beta_i, \tag{20}$$

where

$$\beta_i = \frac{\mathrm{cov}\left(r_{i,t+1}, \xi_{t+1}\right)}{\omega^2}. \tag{21}$$

Figure 1 inverts (3) to show that the risk premium is a convex function of the level of the log Sharpe ratio. The first- and second-order polynomial approximations are

$$\begin{aligned}\widetilde{\Lambda}_t(1) &= \psi_0 + \psi_1 \ln S_t \\ \widetilde{\Lambda}_t(2) &= \psi_0 + \psi_1 \ln S_t + \psi_2 P_2\left(\ln S_t\right),\end{aligned} \tag{22}$$

where ψ_i, $i = 0, 1, 2$, are parameters of the approximation that can be fit by a simple linear regression. $P_2(\cdot)$ was defined earlier. Figure 1 also shows the first- and second-order approximations, $\widetilde{\Lambda}_t(1)$ and $\widetilde{\Lambda}_t(2)$. It is clear that the second-order approximation is virtually exact over the entire reasonable range of $\ln S_t$.

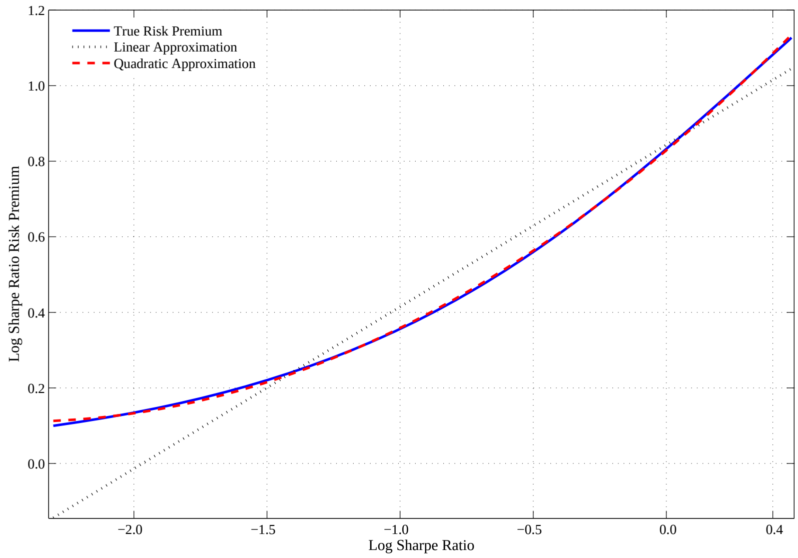

Figure 1: The risk premium on the log Sharpe ratio along with first- and second-order approximations.

The simple reduced-form model described above uses the intuition of a dynamic capital market line to express a pricing kernel that can arise from a variety of underlying structural models. It generates a conditional pricing model that is almost linear. Although there are three state variables that describe the location of the time-varying capital market line, there is only one risk premium driving excess returns. The functional form of this risk premium represents the only source of nonlinearity in the pricing implications of the model. The risk premium is a slightly convex function of the log Sharpe ratio state variable, and it can be well approximated by a second-order polynomial in the log Sharpe ratio.

This structure is ideally suited as an example of the possible pricing implications of misspecifying *modest* nonlinearity in a conditional model. On the one hand, if the model generated substantial nonlinearities, then it would be an example that could easily be dismissed as "rigged" to deliver poor performance of the standard linear approximation. On the other hand, if the model was fully linear, then the standard specification would be completely accurate. The nature and extent of the pricing misspecification introduced by the standard linear approximation approach is examined in the following subsections.

3.2 Estimation

We assume that the econometrician examining data generated by the model in Section 3.1 has access to excess returns and historical values of the market portfolio and the true state variables. The time series regression corresponding to equation (40) is:

$$R^e_{i,t+1} = a_i + b_{i,1}\xi_{t+1} + b_{i,2}\left(\ln s_t\right) + b_{i,3}\left[\xi_{t+1}\cdot\left(\ln s_t\right)\right] + \eta_{i,t+1}, \tag{23}$$

for $i = 1,\ldots,N$. In this case, the true factor is the innovation to the pricing kernel, ξ_{t+1}, the information proxy variable is the lagged value of the log Sharpe ratio, $\ln S_t$, and the conditional component of the model is captured in the interaction term, $\xi_{t+1}\cdot(\ln S_t)$.

Given that $\xi_{t+1} \equiv \delta_1\varepsilon_{1,t+1} + \delta_2\varepsilon_{2,t+1} + \delta_3\varepsilon_{3,t+1}$, (23) expands to

$$\begin{aligned} R^e_{i,t+1} = a_i &+ \delta_1 b_{i,1}\varepsilon_{1,t+1} + \delta_2 b_{i,1}\varepsilon_{2,t+1} + \delta_3 b_{i,1}\varepsilon_{3,t+1} + b_{i,2}\left(\ln S_t\right) \\ &+ \delta_1 b_{i,3}\left[\varepsilon_{1,t+1}\cdot\left(\ln S_t\right)\right] + \delta_2 b_{i,3}\left[\varepsilon_{2,t+1}\cdot\left(\ln S_t\right)\right] \\ &+ \delta_3 b_{i,3}\left[\varepsilon_{3,t+1}\cdot\left(\ln S_t\right)\right] + \eta_{i,t+1}. \end{aligned} \tag{24}$$

Since the δ_i and the original b_i coefficients cannot be identified separately, it is notationally convenient to rewrite (24) as

$$R^e_{i,t+1} = a_i + d_{i,1}\varepsilon_{1,t+1} + d_{i,2}\varepsilon_{2,t+1} + d_{i,3}\varepsilon_{3,t+1} + d_{i,4}(\ln S_t) \tag{25}$$
$$+ d_{i,5}\left[\varepsilon_{1,t+1} \cdot (\ln S_t)\right] + d_{i,6}\left[\varepsilon_{2,t+1} \cdot (\ln S_t)\right] + d_{i,7}\left[\varepsilon_{3,t+1} \cdot (\ln S_t)\right] + \eta_{i,t+1}.$$

The second-pass cross-sectional regression is

$$\overline{R}^e_i = \sum_{j=1}^{7} \lambda_j \widehat{d}_{i,j} + \upsilon_i, \tag{26}$$

for $i = 1, \ldots, N$, with $\widehat{d}_{i,j}$ denoting the fitted value of the coefficients from equation (25) and the fitted value of υ_i corresponding to the measured alpha for asset i.

In the context of this model, the linearity assumption in equation (6) is incorrect. Given the effectiveness of the second-order linear approximation to the true risk premium, a more accurate specification is

$$\begin{aligned} R^e_{i,t+1} &= a_i + c_{i,1}\varepsilon_{1,t+1} + c_{i,2}\varepsilon_{2,t+1} + c_{i,3}\varepsilon_{3,t+1} + c_{i,4}(\ln s_t) + c_{i,5}P_2(\ln S_t) \\ &+ c_{i,6}\left[\varepsilon_{1,t+1} \cdot (\ln S_t)\right] + c_{i,7}\left[\varepsilon_{1,t+1} \cdot P_2(\ln S_t)\right] \\ &+ c_{i,8}\left[\varepsilon_{2,t+1} \cdot (\ln S_t)\right] + c_{i,9}\left[\varepsilon_{2,t+1} \cdot P_2(\ln S_t)\right] \\ &+ c_{i,10}\left[\varepsilon_{3,t+1} \cdot (\ln S_t)\right] + c_{i,11}\left[\varepsilon_{3,t+1} \cdot P_2(\ln S_t)\right] + \eta_{i,t+1}. \end{aligned} \tag{27}$$

The difference between (25) and (27) is the omission of the relevant factors involving $P_2(\ln S_t)$. There is a corresponding (expanded) cross-sectional regression that is analogous to (26).

This comparison makes the source of the misspecification clear. The heart of our analysis is contained in the answers to the following questions: (i) When the true pricing errors (alphas) are zero, what are the magnitudes of the alphas introduced by misspecification of the conditional pricing relation? (ii) How do these true alphas compare with estimates of alphas that might be constructed in a sample of a realistic size drawn from this model? Finally, (iii) does a test statistic of the form in equation (13) for omitted priced factors have any power to detect the omitted variables in a sample of realistic size drawn from this model?

3.3 How Large are the Alphas Caused by Misspecification?

The first step in calibrating the model is to construct proxies for the state variables $\ln S$ and $\ln r^f$. These state variables will be used with a set of asset returns to examine the model's implications. We use the yield on a one-month Treasury bill, from CRSP, as a proxy for the short rate r^f. The sample period is from January 1960 to December 2009 (for a total of $T = 600$ monthly observations). A plot of the log short rate is shown in the top panel of Figure 2. Identifying this state variable is robust to reasonable alternative short rate choices.

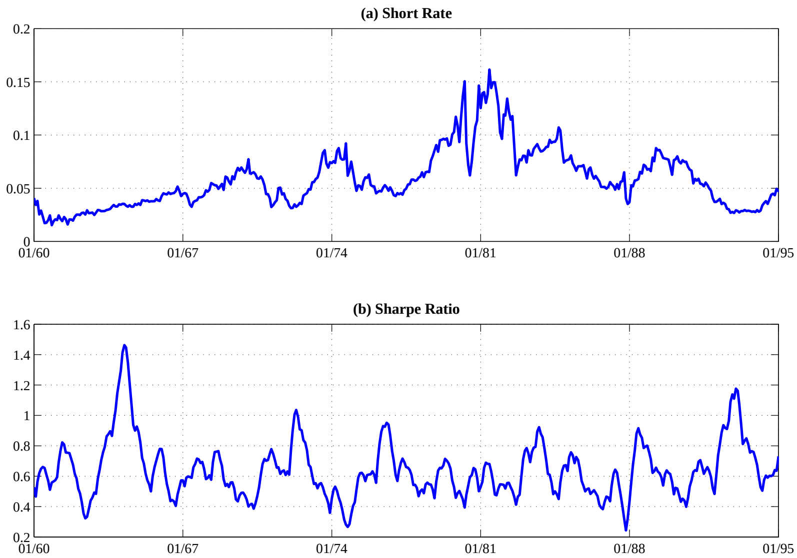

Figure 2: The time series of the pricing kernel state variables.

Identifying the log Sharpe ratio is more difficult. The model in Section 3.1 specifies the state variable as the log Sharpe ratio of the conditionally mean-variance efficient portfolio that prices all traded assets. This portfolio is unobservable, and it is the focus of most modern asset pricing research. Furthermore, even if this asset was observable, its moments would need to be estimated in order to construct the Sharpe ratio. Our approach to calibrating the dynamics of the log Sharpe ratio is to treat the market portfolio as an observable proxy for the conditionally mean-variance efficient portfolio and then to estimate the dynamics of the log Sharpe ratio from historical data on the market portfolio. While this proxy is clearly imperfect, it is not unreasonable to assume that the log Sharpe ratio dynamics of the unobservable conditionally mean-variance efficient portfolio is similar to that of the observable market portfolio, especially for our calibration purposes.

More specifically, we identify the log Sharpe ratio state variable using the filtering approach in Brandt and Kang (2004). They assume that the continuously compounded excess return on the CRSP value-weighted index has both a time-varying conditional mean and a time-varying conditional volatility. The values of these moments are recovered from the realized data using an approximation to the true likelihood based on the Kalman filter. Smoothed estimates of the conditional moments are then constructed based on the full sample of return data. Our estimate of the log Sharpe ratio is constructed from these smoothed estimates. This filtering procedure also recovers a (full sample) estimate of the shock to the market portfolio, ε_3, that can be used in calibrating the pricing kernel.[12] A plot of the log Sharpe ratio, from 1960 to 2009, is shown in the bottom panel of Figure 2.

Panel A of Table 1 presents some simple full-sample summary statistics for the empirical estimates of the two state variables. The continuously compounded return on the Treasury bill averaged 4.9 percent per year over this period with a standard deviation of 2.6 percent per year. It is slightly right-skewed and fat-tailed relative to the normal distribution, and it is highly persistent.[13] The log of the Sharpe ratio averages about -0.57 with a standard deviation of 0.33. The log Sharpe ratio is slightly left-skewed with tails that are as fat as a normal distribution. It is also very persistent but not close to having a unit root.

[12] By construction, these return shocks are normalized to have a standard deviation of one.

[13] The t-test reported in Panel B of Table 4 for the null of $\beta_1 = 1$ is only suggestive. There is a large literature on the failings of standard asymptotic theory in this estimation setting. An augmented Dickey-Fuller test for a unit root in the short rate that includes 12 lagged first differences has a fitted value of -2.513, with a p-value of 0.116. The same test including 18 lagged first differences has a value of -2.629, with a p-value of 0.090.

Table 1: State Variables

Panel A: Summary Statistics					
Series	Mean	Std. Dev.	Skewness	Kurtosis	AR(1)
ln (Short Rate)	0.049	0.026	0.860	4.402	0.975
ln (Sharpe Ratio)	−0.569	0.327	−0.069	2.983	0.941

Panel B: Regression Results				
Series	α_1	β_1		σ_1
	0.001	0.978		
ln (Short Rate)	(0.0006)	(0.013)		0.006
	[18.86]	[−1.746]		
Series	α_2	β_2	γ	σ_2
	−0.017	0.936	−0.381	
ln (Sharpe Ratio)	(0.014)	(0.015)	(0.210)	0.111
	[0.714]	[−4.267]	[−1.814]	

The sample period is from January 1960 to December 2009 ($T = 600$ months). ln (Short Rate) is the continuously compounded return on the one-month Treasury bill. The ln (Sharpe Ratio) is the natural logarithm of the Sharpe ratio extracted from the S&P 500 using the filtering method in Brandt and Kang (2004). Panel A contains simple summary statistics of the full sample. Skewness is the sample estimate of the third central moment standardized by the cube of the sample standard deviation. Kurtosis is the sample estimate of the fourth central moment standardized by the sample standard deviation raised to the fourth power. It is not *excess* kurtosis relative to the normal distribution. AR(1) is the sample estimate of the first-order autocorrelation coefficient. In Panel B, the parameters are estimated using ordinary least squares. The parameter estimates are for equations (15) and (16) in the text. Robust standard error estimates, based on the Newey-West estimator are in parentheses. The brackets are t-tests for $\alpha_i = 0$, $i = 1, 2$, $\beta_i = 1$, for $i = 1, 2$, and $\gamma = 0$. The estimates in this table are constructed using the programs `ReturnMoments.m` and `rp_est.m`. (All of the programs referred to in the notes to this table – and all subsequent tables – are available at http://www2.bc.edu/david-chapman/bcc.html.)

Given the time series of the state variables, we fit equations (15) and (16) to obtain estimates of the parameters that describe these dynamics. Since this system is linear in logs, all of the parameters are estimated by OLS with robust standard errors used for inference. Panel B of Table 1 presents these results. The log Sharpe ratio is significantly positively related to its own lagged value and marginally significantly negatively related to the contemporaneous level of the log short rate. The estimates of the short rate parameters are consistent with the large literature on fitting univariate time series models to short term interest rate data.

The fitted residuals from equations (15) and (16) do not satisfy the model assumption of being *iid* draws from a bivariate standard normal density. This is not surprising, given the evidence in Figure 2 and the extensive literature on the volatility dynamics of both stock returns and short term interest rates. In order to properly calibrate a model with this error structure, we filter the residuals from the regressions in Table 1 using the dynamic conditional correlation multivariate GARCH model of Engle (2002).[14] The transformed shocks correspond more closely to a constant identity matrix covariance matrix for $(\varepsilon_1, \varepsilon_2)'$. They are combined with ε_3 (which is already normalized and orthogonalized with respect to $(\varepsilon_1, \varepsilon_2)'$).

We calibrate asset returns to the 25 portfolios formed by Fama and French (1993) on the basis of market capitalization and book-to-market ratio. The monthly portfolio returns are value-weighted and continuously compounded in excess of the continuously compounded yield on a one-month Treasury bill. The sample period, again, is January 1962 to December 2009. The first four sample moments of excess returns are reported in Table 2. Holding market capitalization constant, the average excess returns increase from low book-to-market to high book-to-market portfolios. In our sample period, there is not a consistent size effect for a given book-to-market rank. Excess return volatility declines with size, and it also declines with the book-to-market ratio but only up to the fourth book-to-market quintile. The returns to all series are negatively skewed and exhibit excess kurtosis (relative to the normal distribution).

[14]This model was fit to the data using Matlab code distributed by the Economics department of the University of California at San Diego.

Table 2: Summary Statistics for the Excess Returns to Size and Book-to-Market Sorted Portfolios

Portfolio	Mean	Std. Dev.	Skewness	Kurtosis
S1 BM1	−0.124	8.211	−0.535	5.650
S1 BM2	0.460	7.011	−0.472	6.279
S1 BM3	0.572	6.072	−0.623	6.454
S1 BM4	0.777	5.724	−0.630	7.025
S1 BM5	0.897	6.163	−0.696	6.949
S2 BM1	0.061	7.483	−0.718	5.395
S2 BM2	0.429	6.140	−0.879	6.748
S2 BM3	0.684	5.539	−0.888	7.065
S2 BM4	0.730	5.372	−0.905	7.106
S2 BM5	0.787	6.119	−0.868	7.105
S3 BM1	0.126	6.881	−0.691	5.185
S3 BM2	0.506	5.563	−0.948	7.099
S3 BM3	0.569	5.086	−0.880	6.245
S3 BM4	0.672	4.931	−0.655	6.217
S3 BM5	0.857	5.549	−0.803	7.143
S4 BM1	0.280	6.062	−0.566	5.154
S4 BM2	0.334	5.313	−1.015	7.676
S4 BM3	0.510	5.119	−0.927	7.714
S4 BM4	0.661	4.877	−0.532	5.331
S4 BM5	0.661	5.559	−0.615	5.905
S5 BM1	0.248	4.833	−0.493	4.881
S5 BM2	0.333	4.544	−0.637	5.473
S5 BM3	0.320	4.437	−0.592	5.930
S5 BM4	0.395	4.403	−0.470	5.307
S5 BM5	0.448	4.970	−0.482	4.597

Returns are continuously compounded in excess of the one-month Treasury bill yield, and they are measured over the period from January 1962 to December 2009. They are reported in percent per month. Skewness is defined as the standardized third central moment of the data, and kurtosis is the standardized fourth central moment (Note: It is not excess kurtosis relative to the normal distribution.). The estimates in this table are constructed using the program `ReturnMoments.m`.

The calibrated values of the alphas generated by the model in Section 3.1 are constructed by computing the required inputs: (*i*) the factor covariance matrices, $\mathbf{\Sigma}_{11}, \mathbf{\Sigma}_{22}$, and $\mathbf{\Sigma}_{12}$; (*ii*) the covariance matrices of the factors with the excess returns, $\mathbf{\Sigma}_{\mathbf{F}_1\mathbf{R}^e}$ and $\mathbf{\Sigma}_{\mathbf{F}_2\mathbf{R}^e}$; and (*iii*) the market risk premium parameters, λ_1 and λ_2. In this setting,

$$\begin{aligned}\underset{(7\times1)}{\mathbf{F}_1} &= \left(\varepsilon_{1,t+1}, \varepsilon_{2,t+1}, \varepsilon_{3,t+1}, \ln S_t, \varepsilon_{1,t+1}\cdot\ln S_t, \varepsilon_{2,t+1}\cdot\ln S_t, \varepsilon_{3,t+1}\cdot\ln S_t\right)' \\ \underset{(4\times1)}{\mathbf{F}_2} &= \left(P_2\left(\ln S_t\right), \varepsilon_{1,t+1}\cdot P_2\left(\ln S_t\right), \varepsilon_{2,t+1}\cdot P_2\left(\ln S_t\right), \varepsilon_{3,t+1}\cdot P_2\left(\ln S_t\right)\right)'\end{aligned} \tag{28}$$

are the included and omitted factors, respectively.

The proxies for the factors are constructed from the fitted (and normalized) shocks to the measured state variables, the shock to the market portfolio, and the lagged levels of the state variables.[15] Given time series estimates for the factors, it is straightforward to construct point estimates of the necessary parameters. The purpose of our estimation is to calibrate the model and not to propose the model as a thorough description of the cross-section of expected returns. The parameters are estimated using Fama-MacBeth with a single time series regression to recover the factor loadings and monthly cross-sectional regressions to estimate the risk premiums and cross-sectional pricing errors. Since the monthly risk premiums and pricing errors are not *iid*, we estimate standard errors using a robust covariance matrix (Newey-West) with an asymptotically optimal bandwidth selection.

Table 3 contains the estimates of the risk premiums from the model that includes second-order approximations in the conditioning information. The innovation to the short rate has a positive risk premium. The other point estimates are predominantly negative. The Shanken-adjusted standard errors are almost twice as large as their uncorrected (but robust) counterparts. There is some evidence consistent with significant risk premiums using only robust standard errors, but this disappears after accounting for measurement error in the factor loadings.

[15] Following the recommendation in Ferson, Sarkissian, and Simin, (2003), we stochastically detrend the (demeaned) conditioning variable, $\ln S_t$ using a 12-month moving average filter. This introduces a two year difference in the starting point for the detrended data.

Table 3: Fitted Risk Premiums January 1962 to December 2009

		Standard Errors	
Term	λ	Robust FM	Robust FM & Shanken
$\varepsilon_{1,t+1}$	0.641	0.373	0.897
$\varepsilon_{2,t+1}$	−0.123	0.286	0.682
$\varepsilon_{3,t+1}$	−0.309	0.463	1.116
$\ln S_t$	0.028	0.241	0.572
$\varepsilon_{1,t+1} \cdot \ln S_t$	−0.393	0.282	0.676
$\varepsilon_{2,t+1} \cdot \ln S_t$	−0.407	0.289	0.696
$\varepsilon_{3,t+1} \cdot \ln S_t$	−0.352	0.406	0.978
$P_2(\ln S_t)$	−0.566	0.302	0.721
$\varepsilon_{1,t+1} \cdot P_2(\ln S_t)$	−0.008	0.182	0.434
$\varepsilon_{2,t+1} \cdot P_2(\ln S_t)$	−0.166	0.213	0.513
$\varepsilon_{3,t+1} \cdot P_2(\ln S_t)$	0.473	0.225	0.542

$\varepsilon_{1,t}$ is the normalized innovation to the short rate at t. $\varepsilon_{2,t}$ is the normalized innovation to the log Sharpe ratio process at t, and $\varepsilon_{3,t}$ is the normalized innovation to the market portfolio at time t. $\ln S_t$ is the log of the Sharpe ratio process at t, and $P_2(\cdot)$ is the orthogonalized and normalized second moment of the relevant process. The estimates of the risk premiums are the sample averages from the second-pass regression

$$\overline{R}_i^e = \sum_{j=1}^{7} \lambda_j \widehat{c}_{i,j} + \upsilon_i,$$

for $i = 1, \ldots, N$, with $\widehat{c}_{i,j}$ denoting the fitted value of the coefficients from time series regression (27). The model is estimated using Fama-MacBeth. The standard error column labeled "Robust FM" uses the Newey-West estimate of the long-run covariance matrix with automatic bandwidth selection. The standard error column labeled "Robust FM & Shanken" adds the Shanken correction term for variability introduced in the first-pass regression to the Newey-West estimator. The estimates in this table are constructed using the program `rp_est_FSS_FM.m`.

The estimates of the alphas (pricing errors) from the full model are shown in Table 4. The model prices 11 portfolios with an average alpha of 10 percent or less of the corresponding average excess return. There are also four portfolios where alphas exceed 30 percent of average excess returns, including the smallest size and B/M portfolio for which the (absolute) alpha is nearly 90 percent of the level of the average excess returns. 10 of the 25 estimated alphas are large relative to the robust standard errors, but none of the alphas are statistically significant after applying the Shanken correction.

Table 4: Fitted Alphas

Portfolio	α	Standard Errors [1]	[2]	Relative α
S1 BM1	-0.109	0.051	0.122	0.880
S1 BM2	0.124	0.066	0.160	0.270
S1 BM3	-0.023	0.035	0.086	-0.041
S1 BM4	0.136	0.075	0.180	0.175
S1 BM5	0.047	0.055	0.134	0.052
S2 BM1	-0.009	0.031	0.076	-0.140
S2 BM2	0.010	0.035	0.083	0.024
S2 BM3	0.097	0.050	0.121	0.142
S2 BM4	0.031	0.042	0.101	0.043
S2 BM5	-0.068	0.039	0.094	-0.087
S3 BM1	-0.111	0.049	0.118	-0.883
S3 BM2	-0.094	0.052	0.127	-0.186
S3 BM3	0.047	0.062	0.149	0.083
S3 BM4	0.003	0.045	0.110	0.004
S3 BM5	0.115	0.041	0.099	0.135
S4 BM1	-0.013	0.058	0.141	-0.047
S4 BM2	-0.074	0.037	0.088	-0.222
S4 BM3	0.017	0.055	0.133	0.033
S4 BM4	0.129	0.056	0.134	0.195
S4 BM5	-0.177	0.078	0.189	-0.268
S5 BM1	0.095	0.046	0.112	0.384
S5 BM2	0.007	0.055	0.133	-0.020
S5 BM3	0.086	0.043	0.105	0.269
S5 BM4	-0.121	0.047	0.113	-0.305
S5 BM5	-0.132	0.052	0.126	-0.294

The data are from January 1962 to December 2009. The model is estimated using Fama-MacBeth with a second-order approximation to the risk premium function; i.e., using equations (25) and (26) in the text. They are reported in percent per month.. The standard error column labeled [1] uses the Newey-West estimate of the long-run covariance matrix with automatic bandwidth selection. The standard error column labeled [2] adds the Shanken correction term for variability introduced in the first-pass regression to the Newey-West estimator. Relative alpha for a given portfolio is defined as the portfolio's alpha divided by the average excess return to the portfolio. The estimates in this table are constructed using the program `rp_est_FSS_FM.m`.

In Table 5, we present estimates of the differences in alphas from the nonlinear and linear models. These differences form the numerator of the quadratic form in the alpha-based nonlinearity test. The final column shows the difference in alphas as a percentage of average absolute excess returns. These differences can be large, both absolutely and relative to average absolute excess returns. The pricing errors for the linear model tend to be larger in the smaller three size quintiles (seen as negative α differences), but the main point is that adding the second-order terms to the pricing kernel has an impact on the measured pricing errors.

Table 5: Change in Alpha Between the Nonlinear and Linear Models

Portfolio	$\Delta\alpha$	Relative $\Delta\alpha$
S1 BM1	0.086	0.694
S1 BM2	−0.013	−0.029
S1 BM3	0.071	0.124
S1 BM4	−0.101	−0.130
S1 BM5	−0.062	−0.069
S2 BM1	−0.000	−0.006
S2 BM2	0.096	0.223
S2 BM3	0.008	0.012
S2 BM4	−0.035	−0.048
S2 BM5	−0.068	−0.086
S3 BM1	−0.088	−0.700
S3 BM2	−0.065	−0.129
S3 BM3	0.050	−0.087
S3 BM4	−0.042	−0.062
S3 BM5	−0.007	−0.009
S4 BM1	−0.014	−0.049
S4 BM2	0.008	0.023
S4 BM3	0.036	0.070
S4 BM4	0.070	0.106
S4 BM5	0.072	0.109
S5 BM1	0.027	0.109
S5 BM2	−0.082	−0.245
S5 BM3	0.063	0.197
S5 BM4	0.016	0.042
S5 BM5	−0.040	−0.090

The data are monthly from Jan. 1962 to Dec. 2009. Alpha changes are quoted in percent per month, and they are constructed by subtracting the alphas from the linear model from the alphas generated by the nonlinear approximation. Relative alphas are defined as the difference divided by the absolute value of the associated average excess returns. The estimates in this table are constructed using the programs `rp_est_FSS_FM.m` and `rp_est_fss_lin.m`.

3.4 A Monte Carlo Study

The primary reason for considering a Monte Carlo study at this point is the issue of detecting the presence of nonlinearity in the risk premium or, equivalently (in this setting), the coefficients of the pricing kernel. Do common tests of the form of (12) or (13), described in Section 2, have reasonable size and power against a specific local alternative? If not, then the only symptom of nonlinearity are the measured pricing errors, which generally reflect a wide range of possible model misspecification, not only nonlinearity of the risk premiums.

The general structure of the experiment we consider is:

Step #1: Generate $5,000$ independent simulated sample paths of an economy in which returns are generated according to the model economy calibrated to the point estimates of size and B/M sorted returns. Excess returns on the different portfolios are constructed by:

Step #1a: Simulating values for the model state variables, $(\ln S_t, \ln r_t^f)'$, using the estimated parameter values in Table 4 and bivariate standard normal shocks.

Step #1b: Along each simulated sample path, realized excess returns are generated as the sum of expected returns and unexpected returns. Time-varying expected returns are generated according the model, the calibrated parameter values, and the realized values of the state variables generated in Step #1a; i.e., using (20) and (22):

$$E_t(r_{i,t+1}) + \frac{1}{2}\text{var}_t(r_{i,t+1}) - r_t^f = [\psi_0 + \psi_1 \ln S_t + \psi_2 P_2(\ln S_t)]\,\beta_i,$$

for each i in N, where β_i is defined in equation (18). Unexpected returns are generated at each date in the simulation by drawing from a multivariate normal distribution with mean zero and a diagonal covariance matrix with the nonzero elements of the matrix set equal to the estimates of the diagonal elements of the covariance matrix of the shocks from equation (23).

Step #1c: Along each simulated path of the state variables, we generate returns by a separate linear version of the model economy in order to assess the actual size of the nonlinearity tests.

Step #2: Each simulation is of length 600 months, and the state variables and realized excess returns are collected after a "burn-in" period of 4,900 months designed to reduce the influence of initial conditions.[16]

Step #3: For each simulated path:

Step #3a: we estimate a standard linearized version of the conditional model, as in equations (23) and (26). The model is estimated using a standard Fama-MacBeth

[16] The unconditional means of the state variables are used as the initial conditions for each simulation. In order to mitigate problems caused by the near unit root properties of the estimated short rate, we discard sample paths whose first value after the burn-in period is more than three standard deviations from the unconditional mean of the short rate. This adjustment is important because the nonlinearity becomes more important in the tails of the distribution, and it is therefore designed to make our results conservative.

algorithm. The alphas (pricing errors) for each asset and the estimated risk premiums in each simulation are stored for later examination.

Step #3b: We also estimate a version of an approximate model that includes second-order terms in the conditioning variables. We construct the test statistics in (12) and (13) – both with and without a correction for estimation of the betas in the first-pass time series regression – for the null hypothesis that the higher-order terms are all jointly equal to zero. We store these results for later examination (of the power of the tests against the specific alternative).

Step #4: Repeat Steps (1) through (3) using more nonlinearity in the pricing kernel by slightly adjusting the higher-order parameter values to generate a slighlty larger spread in the simulated test asset returns.

The first column of Table 6 shows the average excess returns of the test portfolios in the baseline parameterization. These are the cross-simulation averages of the time series average excess returns along each simulated sample path. The standard deviation of excess returns reported in the second column is the cross-simulation average of the time series standard deviation computed along each simulated path. The simulated excess returns of the characteristics sorted portfolios in Table 6 are generally reasonable, with the exception of the negative excess return on the portfolio formed from the smallest size and lowest B/M stocks. For the most part, they are comparable in magnitude to the actual returns reported in Table 2 (albeit somewhat larger), and there is a value premium. However, as with the actual data over this sample period, there is little or no evidence of a consistent size effect. The second pair of columns shows average excess returns for the more nonlinear simulation.

Table 6: Excess Returns in the Simulated Example Economies

	Baseline		More Nonlinear	
Portfolio	Mean	StdDev	Mean	StdDev
S1 BM1	−0.026	8.381	−0.926	8.382
S1 BM2	0.343	6.793	0.126	6.793
S1 BM3	0.593	5.978	0.509	5.979
S1 BM4	0.640	5.666	0.866	5.666
S1 BM5	0.844	6.167	1.179	6.167
S2 BM1	0.076	7.110	−0.587	7.110
S2 BM2	0.421	5.718	0.366	5.718
S2 BM3	0.588	5.291	0.602	5.291
S2 BM4	0.699	5.022	1.064	5.022
S2 BM5	0.849	5.402	1.729	5.402
S3 BM1	0.245	6.532	−0.188	6.532
S3 BM2	0.604	5.414	0.686	5.414
S3 BM3	0.527	4.848	0.721	4.848
S3 BM4	0.675	4.916	1.148	4.916
S3 BM5	0.740	5.110	1.281	5.110
S4 BM1	0.303	5.924	−0.047	5.924
S4 BM2	0.412	5.258	0.309	5.258
S4 BM3	0.498	4.961	0.481	4.961
S4 BM4	0.530	4.787	0.651	4.787
S4 BM5	0.834	5.303	1.406	5.303
S5 BM1	0.163	4.995	−0.246	4.995
S5 BM2	0.346	4.361	0.229	4.361
S5 BM3	0.236	4.300	0.203	4.300
S5 BM4	0.516	4.459	0.851	4.459
S5 BM5	0.580	4.912	0.849	4.912

This table reports average excess returns, average standard deviation of excess returns, and alphas – both absolute and relative – from $5,000$ simulations of length $T = 600$ months of the model economy in Section 3.1 calibrated to size and book-to-market sorted portfolio returns. Each entry corresponds to the cross-simulation average of the time series average of each component of excess returns or alphas. Alphas are constructed from the application of the linear approximation to the two parameterizations of the nonlinear model. The baseline parameterization uses the coefficient estimates from Table 1 to estimate the nonlinear factors. The more nonlinear parameterization provides a slight increase in the nonlinear risk premium parameters. OLS is used in fitting the second-pass cross-sectional regression. True alphas are set to zero in the simulations. These estimates are constructed in the programs `Sim01_fss_FM.m` and `Sim02_fss_FM.m`.

Table 7 shows the alphas and relative alphas from both parameterizations of the model. By construction, the true alphas are zero in the correctly specified model. Again, these are the cross-sectional averages of the time series average from each of the $5,000$ simulations. In the baseline simulation, the average alphas are largest (in absolute value) for the lowest B/M portfolio in four of the five size quintiles. The largest (in absolute value) alphas are around 0.25 to 0.30 percent per month (3.0 to 3.6 percent per year). Relative to average expected returns, the largest alphas are between (roughly) 25 percent to 152 percent of average excess returns. The cross-sectional standard deviation of average alphas across the different portfolios ranges between 0.11 and 0.14 percent per month. Since these are the cross-simulation moments, the mean and standard deviations indicate that a substantial number of the $5,000$ simulations generated large alphas. Relative alphas are more variable than alphas, indicating that there are many sample paths where estimated alphas are large relative to average excess returns.[17] For the lowest B/M assets in the two smallest size quintile, the standard deviations of the distributions of relative pricing errors are quite large. The average alphas in the more nonlinear parameterization are small for the smaller two size quintiles but increase substantially in both absolute and relative amounts for the larger size quintiles.

[17]There are some simulated assets with low levels of average returns; e.g., small size and low B/M (see Table 6). In these cases, the denominator of the relative alpha may be close to zero in some simulated return series, and this has the effect of "blowing up" even a relative small estimated alpha. As a result, we do not emphasize the relative alpha results for the 6 assets with the lowest average excess return in Table 6.

Table 7: Simulation Results for α in the Example Economies

	Baseline				More Nonlinear			
	α		Relative α		α		Relative α	
Portfolio	Mean	StdDev	Mean	StdDev	Mean	StdDev	Mean	StdDev
S1 BM1	−0.297	0.137	7.488	429.3	−0.814	0.476	1.202	8.832
S1 BM2	−0.113	0.124	−0.359	0.406	−0.293	0.344	−0.299	159.7
S1 BM3	0.047	0.120	0.074	0.202	−0.064	0.323	−0.147	0.698
S1 BM4	0.080	0.119	0.123	0.185	0.180	0.323	0.201	0.345
S1 BM5	0.203	0.127	0.239	0.148	0.372	0.351	0.309	0.267
S2 BM1	−0.260	0.132	−0.222	117.0	−0.699	0.439	1.277	52.48
S2 BM2	−0.063	0.116	−0.161	0.288	−0.152	0.315	−0.453	0.965
S2 BM3	0.062	0.115	0.103	0.196	0.026	0.308	0.040	0.509
S2 BM4	0.143	0.115	0.202	0.162	0.362	0.330	0.329	0.272
S2 BM5	0.226	0.120	0.265	0.138	0.798	0.465	0.442	0.187
S3 BM1	−0.159	0.128	−0.734	0.685	−0.473	0.390	4.048	315.2
S3 BM2	0.079	0.114	0.127	0.189	0.091	0.305	0.130	0.432
S3 BM3	0.032	0.109	0.058	0.208	0.140	0.301	0.1840	0.387
S3 BM4	0.132	0.111	0.193	0.163	0.439	0.347	0.366	0.251
S3 BM5	0.178	0.115	0.239	0.152	0.529	0.367	0.397	0.231
S4 BM1	−0.121	0.119	−0.433	0.445	−0.394	0.356	2.013	77.13
S4 BM2	−0.047	0.112	−0.124	0.281	−0.156	0.305	−0.611	1.293
S4 BM3	0.005	0.109	0.005	0.2221	−0.049	0.292	−0.109	0.623
S4 BM4	0.023	0.109	0.040	0.207	0.068	0.292	0.098	0.427
S4 BM5	0.224	0.121	0.267	0.142	0.579	0.384	0.398	0.219
S5 BM1	−0.219	0.119	−1.636	7.187	−0.549	0.371	1.466	48.67
S5 BM2	−0.096	0.106	−0.289	0.323	−0.221	0.293	−1.281	2.469
S5 BM3	−0.174	0.110	−0.775	0.537	−0.224	0.296	−1.201	1.760
S5 BM4	0.033	0.108	0.059	0.210	0.257	0.320	0.279	0.329
S5 BM5	0.082	0.110	0.139	0.190	0.249	0.317	0.281	0.334

This table reports average excess returns, average standard deviation of excess returns, and alphas – both absolute and relative – from $5,000$ simulations of length $T = 600$ months of the model economy in Section 3.1 calibrated to size and book-to-market sorted portfolio returns. Each entry corresponds to the cross-simulation average of the time series average of each component of excess returns or alphas. The baseline parameterization uses the coefficient estimates from Table 1 to estimate the nonlinear factors. The more nonlinear parameterization provides a slight increase in the nonlinear risk premium parameters. OLS is used in fitting the second-pass cross-sectional regression. True alphas are zero. Misspecification alphas (based on the true moments of the data) are shown in the first column of Table 4. These estimates are constructed in the programs `Sim01_fss_FM.m` and `Sim02_fss_FM.m`.

Is it possible to detect the nonlinear components of risk premiums using the nonlinearity test statistics in (12) and (13), respectively? Before we consider the power of the test against the null of modest nonlinearity in the risk premiums, it is important to establish the actual size of the different test statistics. In order to do this, we calibrate a fully linear version of the model in Section 3.1, one that conforms to the null of zero coefficients in higher-order risk premium terms. We then simulate $5,000$ sample paths of length $T_{sim} = 600$ months and compute the different versions of the test statistic along each sample path. The resulting simulated distributions are then stored for later examination.

Table 8 contains estimates of the actual size of the nonlinearity tests from these simulations. Panel A examines the risk premium test (λ-test), estimated using risk premium covariances based on both simple time series estimates, with and without a Shanken correction, and methods that are robust to serial correlation and heteroskedasticity in the measured risk premiums, again, with and without a Shanken correction.[18] Panel B (Panel C) contains the same calculations for the α-test (KRS test).

As Panel A of Table 8 shows, using either a simple or a robust estimate of the covariance matrix, the uncorrected tests both have a smaller right tail than the corresponding asymptotic distribution resulting in too few rejections of the true null at either the 5% or 1% nominal levels. At the 10% nominal level, the test statistic based on the simple (robust) covariance matrix estimate slightly over- (under-) rejects. The actual size of the Shanken-corrected risk premium tests seems to be quite distorted relative to the asymptotic distribution. The apparent overcorrection in the estimated covariance matrix may follow from the simple iid nature of volatility in the example economy.

In the α-tests examined in Panel B of Table 8, the results are quite different. The uncorrected test statistics, whether constructed with a simple or robust covariance matrix estimate, significantly over-reject the true null at the nominal 10% and 5% levels, and the statistic based on the robust estimator continues to over-reject at the 1% level. The Shanken-correction appears to be very important here in getting the pricing error covariance matrix approximately correct since the size distortion is much lower using the Shanken-corrected estimators. Panel C of Table 8 shows that the GLS version of the KRS has reasonable finite-sample size.

[18]The robust covariance matrix estimates us the Newey-West estimator with the asymptotically optimal lag length choice.

Table 8: Size of the Nonlinearity Tests Based on Monte Carlo Simulations

Panel A: Risk Premium Test (λ-test)			
	Nominal Size		
	0.10	0.05	0.01
Simple VCV w/o Shanken Correction	0.124	0.034	0.001
Robust VCV w/o Shanken Correction	0.070	0.016	0.001
Simple VCV w/ Shanken Correction	0.000	0.000	0.000
Robust VCV w/ Shanken Correction	0.000	0.000	0.000
Panel B: Alpha Test (α-test)			
	Nominal Size		
	0.10	0.05	0.01
Simple VCV w/o Shanken Correction	0.671	0.315	0.007
Robust VCV w/o Shanken Correction	0.810	0.549	0.151
Simple VCV w/ Shanken Correction	0.100	0.018	0.000
Robust VCV w/ Shanken Correction	0.223	0.103	0.020
Panel C: *KRS*-Test			
	Nominal Size		
	0.10	0.05	0.01
OLS	0.001	0.000	0.000
GLS	0.135	0.071	0.002

"Simple VCV" refers to the variance-covariance matrix (of either risk premiums or alphas) estimated from the time series of the cross-sectional averages without any correction for time-variation or heteroskedasticity. "Robust VCV" refers to the variance-covariance matrix (of either risk premiums or alphas) estimated from the time series of the cross-sectional averages based on the Newey-West long-run variance-covariance matrix estimator using the asymptotically optimal lag length parameter. The risk premium based test in Panel A is defined in equation (13) in the text, and it has an asymptotic χ^2_4 distribution. The alpha based test in Panel B is defined in equation (12) in the text, and it also has an asymptotic χ^2_4 distribution. The KRS-test in Panel C is defined in equation (14) in the text. The asymptotic critical values are tabulated in the program `cvalcalc.m`. The Shanken-correction to the estimated variance-covariance matrices is defined in Appendix B. The table reports the actual size (i.e., the percentiles of the empirical distribution), based on the critical regions defined by these critical values, in $5,000$ simulated sample paths of the model economy in Section 3.1. These estimates are all constructed in the program `simtestsize_new.m`.

Given our findings in Table 8, it makes sense to examine the power of only the versions of the tests that have reasonable size properties. Furthermore, we do not rely on the critical values for the tests based on the asymptotic χ^2-distribution. Instead, we use the simulated-based critical values values tabulated in the construction of Table 8. We present these results in Table 9. Panel A presents the results from the baseline nonlinear model parameterization, while Panel B shows the power of the tests to distinguish the linear from the more nonlinear parameterization. The the λ-test is more powerful at detecting modest nonlinearity in risk premiums. For example, when the empirical size is set to 0.10, the λ-tests rejects the null of linearity roughly 27 percent of the time in the baseline parameterization. This rate jumps up to almost 90 percent in the more nonlinear parameterization (which gives us a second point on the power curve). In contrast, the α-tests never reject the false null of linearity in the baseline parameterization, although they do have similar power to the λ-tests in the (slightly) more nonlinear parameterization. These findings seem reasonable since pricing errors reflect all forms of misspecification in the model whereas the λ-tests are designed to focus on the magnitudes of the risk premium terms.

Table 9: Power of the Nonlinearity Tests Based on Monte Carlo Simulations

Panel A: Simulations of the Baseline Economy			
	Empirical Power		
	0.10	0.05	0.01
λ-test: Simple VCV w/o Shanken Correction	0.273	0.198	0.057
α-test: Simple VCV w/o Shanken Correction	0.000	0.000	0.000
α-test: Robust VCV w/ Shanken Correction	0.006	0.004	0.002
KRS-Test (GLS)	0.112	0.057	0.013
Panel B: Simulations of the More Nonlinear Economy			
	Empirical Power		
	0.10	0.05	0.01
λ-test: Simple VCV w/o Shanken Correction	0.895	0.875	0.825
α-test: Simple VCV w/o Shanken Correction	0.846	0.832	0.811
α-test: Robust VCV w/ Shanken Correction	0.132	0.103	0.057
KRS-Test (GLS)	0.110	0.058	0.013

The λ-test (α-test) is defined in equation (13) (equation (12)). "Simple" versus "Robust" means that the variance-covariance matrix in the test statistic is estimated as the simple covariance from the time-series of Fama-MacBeth estimates or the Newey-West robust variance covariance matrix, respectively. The Shanken-correction to the estimated variance-covariance matrices is defined in Appendix B. The *KRS*-tests are defined in equation (14) in the text. As documented in Table 8, these versions of the test statistics suffer from size distortion (although substantially less distortion than the risk premium test). The *empirical critical values* associated with the different versions of the test statistics are:

Empirical Size	0.100	0.050	0.001
λ-test: Simple VCV w/o Shanken Correction	6.206	6.986	8.489
α-test: Simple VCV w/o Shanken Correction	18.635	19.803	21.896
α-test: Robust VCV w/ Shanken Correction	17.058	18.829	23.268
KRS-Test (GLS)	63.698	88.438	148.29

The table reports the actual rejection rates based on the critical regions defined by these critical values, in $5,000$ simulated sample paths of the linear version of the model economy in Section 3.1. These estimates are constructed in the programs `Sim_Lin_fss_FM.m` and `simtestpower.m`

4 Application: A Model of Technological Innovation

This example is based on the general equilibrium production economy in Papanikolaou (2011) as extended by Kogan, Papanikolaou, and Stoffman (2012) (hereafter KPS).[19] The model has implications for the role of technological innovation in determining different moments of asset returns, including the value premium. The model implies that the pricing kernel depends on innovations to both disembodied and embodied (in specific vintage capital) productivity shocks. The model is formulated in continuous-time and the SDF evolves as

$$\frac{d\pi_t}{\pi_t} = -r_{f,t}dt - \gamma_x(\omega_t)\, dB_t^x - \gamma_\xi(\omega_t)\, dB_t^\xi, \tag{29}$$

where $r_{f,t}$ is the instantaneous risk-free rate, dB_t^x is the innovation to disembodied (labor-augmenting) productivity shock, and dB_t^ξ is the innovation to the embodied technology shock.[20] Given (29), the instantaneous excess return to any asset i is

$$E_t\left[\frac{dS_{i,t} + D_{i,t}dt}{S_{i,t}} - r_{f,t}dt\right] = \gamma_x(\omega_t)\operatorname{cov}_t\left(\frac{dS_{i,t}}{S_{i,t}}, dB_t^x\right) + \gamma_\xi(\omega_t)\operatorname{cov}_t\left(\frac{dS_{i,t}}{S_{i,t}}, dB_t^\xi\right), \tag{30}$$

where $S_{i,t}$ is the ex-dividend price of asset i and $D_{i,t}$ is the instantaneous dividend payout rate for asset i. This has the standard interpretation of decomposing the excess return into prices of risk ($\gamma_x(\omega_t)$ and $\gamma_\xi(\omega_t)$) and measures of risk, defined as the (conditional) covariances of instantaneous ex-dividend returns with the innovations to the two shocks.

The state variable determining (in part) the market prices of risk is ω_t, a linear function of the two technology shocks and the log-level of physical capital in the economy. These prices do not have closed-form solutions, but for the model parameterization examined in KPS, they look like:

[19] See also the related analysis in Kogan and Papanikolaou (2012).
[20] Equation (29) is equation (40) in KPS.

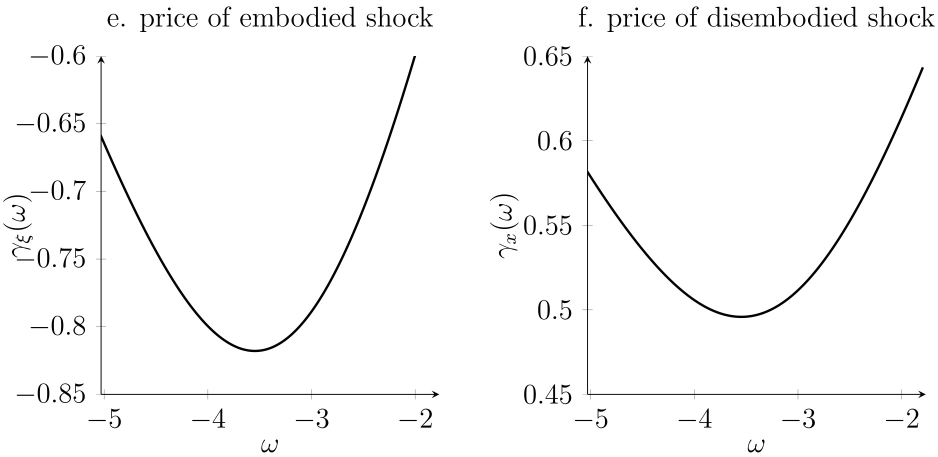

Figure 3: Risk Premium Functions in the KPS Model.

where the figure is reproduced from a portion of Figure 3 in KPS. As KPS note, "(t)he effect of the embodied shock is (generally) ambiguous; following a positive embodied shock, shareholders capture a smaller slice of a larger pie. The net effect ... depends on preference parameters .... The conditional market price of the disembodied shock is positive .... (T)he key difference in comparison to the embodied shock is that a positive disembodied shock also has a large positive effect on the consumption of stockholders ... (through) increased labor productivity."[21]

The interesting point for our purposes is the nonlinear and conditional nature of the model implied risk premiums. The choice of scale in a figure is always (somewhat) discretionary, and so what appears to be substantial nonlinearity in the risk premiums may, in fact, be quite modest in the data given the noisy nature of expected returns. In their analysis, KPS choose to focus on the unconditional pricing implications of a linearized version of the SDF.[22]

We will actually estimate the simpler, but qualitatively similar, model from Papanikolaou (2011) where the the pricing equation analogous to equation (30) is

$$E_t\left[\frac{dS_{i,t}+D_{i,t}dt}{S_{i,t}}-r_{f,t}dt\right]=\gamma_x\mathrm{cov}_t\left(\frac{dS_{i,t}}{S_{i,t}},dB_t^x\right)+\gamma_\xi\left(\omega_t\right)\mathrm{cov}_t\left(\frac{dS_{i,t}}{S_{i,t}},dB_t^\xi\right), \tag{31}$$

where γ_x is the constant price of market-wide risk affecting all firms and $\gamma_\xi\left(\omega_t\right)$ is the nonlinear price of embodied ("investment-specific") technology shocks. In this case, the state variable, ω_t, is correlated with the economy's log investment-to-consumption ratio; so we use this as an

[21] See KPS pages 22-23.

[22] See Section 5.4 – in particular eq. (52) – and Table 9 in KPS. Similar econometric tests are conducted in Papanikolaou (2011) and Kogan and Papanikolaou (2012).

empirical proxy in our conditional estimation.[23]

The "beta" version of the extended security market line in equation (31) is

$$E_t\left[R_{i,t+1}\right] = \Lambda_{0,t} + \Lambda_1 \beta^x_{i,t} + \Lambda_{2,t}\beta^{\xi}_{i,t}, \tag{32}$$

where $R_{i,t+1}$ is the excess return to one of N test assets measured from t to $t+1$, $\Lambda_{0,t}$ is a (possibly) time-varying intercept (reflecting possible measurement error in the risk-free rate proxy), Λ_1 is a scalar (constant) risk premium on market-wide risk, $\Lambda_{2,t}$ is the conditional risk premium on embodied technological change, and $\beta^j_{i,t}$ for $j \in \{x, \xi\}$ is the conditional beta of asset i with respect to a source of risk.

If we apply the standard approximations to estimate the conditional model

$$\begin{aligned} \Lambda_{0,t} &= a_0 + a_1 \Delta ic_t \\ \Lambda_1 &= b_0 \\ \Lambda_{2,t} &= c_0 + c_1 \Delta ic_t, \end{aligned} \tag{33}$$

where $\Delta ic_t \equiv \ln\left({}^{I_t}/_{C_t}\right)$, and the law of iterated expectations, the expanded unconditional pricing model is

$$E\left[R_{i,t+1}\right] = \lambda_0 + \lambda' \beta_i. \tag{34}$$

β_i are the betas of asset i with respect to the expanded set of factors formed from interacting the constant and the consumption growth factors in (32) with the conditioning variable Δic.[24] These factors can be denoted as

$$\underset{(5\times 1)}{\mathbf{F}^{TC}_{t+1}} = \left(1, \Delta ic_t, \Delta c_{t+1}, \xi_{t+1}, \Delta ic_t \cdot \xi_{t+1}\right)', \tag{35}$$

where the change in the aggregate consumption growth, Δc_{t+1}, is the proxy for the market-wide risk, ξ_{t+1} is a proxy for technological innovation based on the spread in the equity returns

[23] Following Kogan and Papanikolaou (2012), we use the log ratio of (real) non-residential private investment to (real) consumption of nondurables plus services; see the discussion on page 16 of Kogan and Papanikolaou (2012).

[24] Equation (31) also suggests that factor loadings are time-varying. we would want to account for this in a general estimation of the technological innovations model. However, in the interest of focusing on the importance of nonlinear conditional risk premiums, we have assumed constant betas in our conditional model.

between investment and consumption sector firms.[25] In constructing (35), we impose the model-implied constraint that the market-wide risk premium is constant. There is no simple connection between the true conditional risk premium, $\Lambda_{2,t}$, and the (relevant) coefficients λ in (34). The computation of $\Lambda_{2,t}$ from λ requires knowledge of the conditional covariance of the expanded factors.[26] In order to illustrate the impact of potential nonlinearities in risk premiums, we consider an additional set of factors

$$\underset{(2\times1)}{\widetilde{\mathbf{F}}_{t+1}^{TC}} = \left(P_2\left(\Delta ic_t\right), P_2\left(\Delta ic_t\right)\cdot\xi_{t+1},\right)', \tag{36}$$

where $P_2\left(\Delta ic_t\right)$ is the component of Δic_t^2 that is orthogonal to Δic_t, normalized to have zero mean and the same unconditional volatility as Δic_t.

The results of estimating the technological change model model, using both $\mathbf{F}^{TC}$ and an extended model using $\left(\mathbf{F}^{TC\prime}, \widetilde{\mathbf{F}}^{TC\prime}\right)'$, are shown in Table 10. The data consist of $T = 259$ quarters of data, from the second quarter of 1948 to the fourth quarter of 2012.[27] The test assets consist of the 25 size- and B/M-sorted portfolios that have become the standard in empirical work.

[25] See Appendix C.

[26] See, for example, the discussion in Lettau and Ludvigson (2001).

[27] Given the size of our cross-section of test assets, we need as long a time series of returns and factors as possible; see the simulation results in Ferson and Foerster (1994).

Table 10: Estimates of the Technological Innovation Model

Panel A: Parameter Estimates

	Linear		Nonlinear	
	Value	"*t-stat*"	Value	"*t-stat*"
Constant	0.0156	2.355	0.0280	5.0315
Δic_t	−0.0181	−0.490	−0.0353	−0.9914
Δc_{t+1}	−0.0006	−0.381	−0.0014	−1.1780
ξ_{t+1}	−0.0137	−1.464	−0.0259	−3.1508
$\Delta ic_t \cdot \xi_{t+1}$	0.0161	0.973	0.0381	2.8105
$P_2(\Delta ic_t)$	n.a.	n.a.	−0.7240	−2.2066
$P_2(\Delta ic_t) \cdot \xi_{t+1}$	n.a.	n.a.	−0.0337	−2.1456

Panel B: Nonlinearity Tests

Test Statistic	Value	Asymptotic P-Value	Empirical P-Value
λ-test: Simple VCV	8.579	0.014	0.609
α-test: Simple VCV	9.000	0.011	0.590

Panel A of the table reports both the estimates of λ from a linear conditioning version of the technological innovation model and from the expanded model that includes nonlinear risk premium terms estimated using the Fama-MacBeth. The data is from 1948:Q2 to 2012:Q4. "*t-stat*" is the ratio of the point estimate to the Fama-MacBeth standard error estimated from the quarterly coefficient estimates allowing for time-variation in the coefficients through a robust variance-covariance estimator. $P_2(\cdot)$ is the function that orthogonalizes $(\Delta ic_t)^2$ with respect to Δic_t and normalizes the orthogonalized variable to have the same magnitude as Δic_t. The nonlinearity tests in Panel B are defined in equations (13) and (12) in the text they are implemented without the Shanken correction for measurement error in the first-pass beta estimates. The asymptotic P-values are based on χ^2_2 distribution and the empirical P-values are based on 5,000 simulations from a block bootstrap of realized returns and factors. These estimates are constructed in the programs `TCmodelest.m` and `TCmodeltTest.m`.

In addition to the point estimates of the risk premiums, λ, we report "t-statistics" formed from the ratio of the estimated coefficient and the coefficient's standard error based on a Fama-MacBeth procedure using a robust covariance matrix estimator that is robust to deviations from the iid assumption in the basic form of Fama-MacBeth. The results in Table 10 are consistent with our basic point about the importance of explicitly controlling for modest nonlinearity in conditional risk premiums. In the first two columns of Panel A, there is little evidence of cross-sectional pricing ability of the model with only linear conditioning. The point estimates of the prices of risk are small, both economically and in relation to their standard errors.[28]

Adding the nonlinear conditioning in the risk premiums, shown in the second pair of columns in Panel A of Table 10, has a significant impact on the estimated coefficients. In particular the risk premium for the embodied technological change factor, ξ_{t+1}, is twice as large and quite large relative to its standard error (as shown by the relevant entry in the t-statistic column). The coefficient estimates on the higher-order conditioning terms are large relative to their standard errors. Moreover, $P_2\left(\Delta ic_t\right) \cdot \xi_{t+1}$, suggests a *concave* rather than the convex risk premium on the embodied technological change factor found in the earlier work.

Although the individual higher order terms appear to be important, the formal nonlinearity tests – when adjusted for finite-sample distributions – do not suggest that the nonlinear terms are statistically significant. The individual pricing errors for the 25 test assets are shown in Figure 4 for both the linear and nonlinear versions of the model. The nonlinear model appears to offer a modest improvement in the fit of the model.

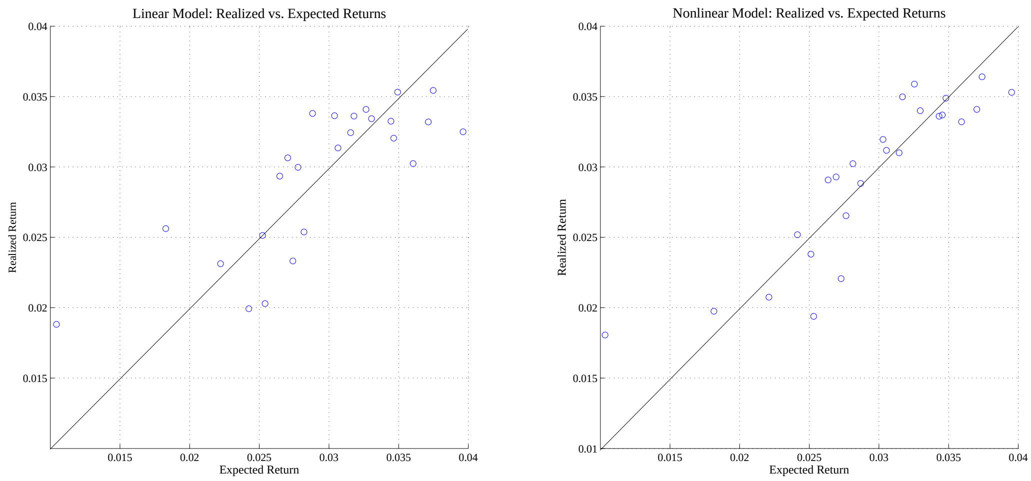

Figure 4: Realized vs. Expected Returns in the Linear and Nonlinear IST Shocks Model. The data are quarterly from the second quarter of 1948 to the fourth quarter of 2012. Each

[28]Papanikolaou (2011) estimates a linear unconditional version of his model using a variety of test assets and proxies for investment-specific technology shocks. In the single version of his (many) tests that is closest to the findings in Table 10, there is only very marginal statistical significance for the linear IST shock factor. We note, as well, that our estimation period includes both prior (15 years) and subsequent (4 years) data when compared to Papanikolaou (2011). This extended time period required us to construct our own measure of the investment shock. A complete description of the construction of this series and a comparison with the series from Papanikolaou (2011) is contained in Appendix C. We would like to thank Dimitris Papanikolaou for making his spread portfolio return series available to us

circle (in each plot) corresponds to one of the 25 size- and B/M-sorted portfolios used as test assets. This picture is constructed in the program `TCmodelest.m`.

5 Conclusions

If conditional risk premiums are nonlinear, what happens when a standard linear approximation is used to estimate the model? This question is important because theoretical models rarely provide guidance to an econometrician on how to handle the unobservable dynamics of conditioning information. We have shown, under a general approximation to a nonlinear model, that using a linear approximation introduces an omitted variables bias associated with the higher order terms in the nonlinear approximation. We derived the forms of the pricing error bias, and we examined alternative forms of specification tests for nonlinearity.

We conduct a Monte Carlo study to quantify the possible importance of *modest* nonlinearity on estimation results based on the assumption of linear conditioning. Our findings suggest that, despite the general noise in estimating expected returns, misspecifying the form of the conditional risk premium can contribute significantly to measured pricing errors. The Monte Carlo results also provide guidance as to the form of the nonlinearity test that is most likely to be useful in practice: a direct test of nonlinear risk premiums using an explicit correction for errors-in-variables in estimating the factor loadings in the first-pass time series regression.

Finally, we examine the implications for misspecifying the risk premium in an example of a recent structural asset pricing. Our empirical analysis of the embodied technological change model of Papanikolaou (2011) clearly demonstrates the importance of allowing for nonlinearity in evaluating the model's performance.

A The "Standard" Testing Approach

The standard testing approach consists of four steps:

Step #1: Start with the fundamental asset pricing equation:

$$E\left(M_{t+1}\mathbf{R}^e_{t+1} \mid \mathcal{F}_t\right) = \mathbf{0}, \tag{37}$$

where $\mathbf{R}^e_{t+1}$ is an N-vector of marketed asset returns from t to $t+1$ in excess of a risk-free return proxy, $\mathbf{0}$ is an N-vector of ones, and $\mathcal{F}_t$ denotes the information used by the market in determining prices at time t.

Step #2: Use a linear structure for the asset pricing kernel:

$$M_{t+1} = 1 + \mathbf{f}'_{t+1}\mathbf{b}_t, \tag{38}$$

as either an exact representation of the model being tested (including possibly a log transformation) or as an approximation to the true model. The use of excess returns results in normalizing the conditional mean of the pricing kernel to 1. M_{t+1} is a conditional pricing model if the vector of pricing kernel coefficients, $\mathbf{b}_t$, are time-varying with the information in $\mathcal{F}_t$.

The essence of the linear conditioning approximation is contained in the following step:

Step #3: Reduce the information set, $\mathcal{F}_t$, to a K-vector of observable variables, $\mathbf{Z}_t$, and write the pricing kernel coefficient, $\mathbf{b}_t$, as

$$\mathbf{b}_t = \mathcal{B}\mathbf{Z}_t, \tag{39}$$

where $\mathcal{B}$ is a $K \times L$ matrix of constant coefficients.[29]

It is important to note that the linearity assumption in (39) is distinct from the linear form of the pricing kernel. There is often no obvious reason to impose linearity in the true model state

[29] If the first element of $\mathbf{Z}$ is $\mathbf{1}$, then there is an intercept in the pricing kernel. As Cochrane (2005) notes, if $\mathbf{Z}_t$ contains all measurable transformations of all elements of the random vector whose minimal σ-algebra defines $\mathcal{F}_t$, then this is not an assumption but rather a mathematical fact. The *assumption* is that a specific finite set of $\mathbf{Z}$ is an adequate approximation to this larger set.

variables – much less linearity in the proxies $\mathbf{Z}$ – on the pricing kernel coefficients. The use of the linearity assumption is a direct result of the econometrician's inability to specify a complete structure for the dynamics of investors' information sets.

The implications of the fundamental asset pricing equation (37) have now been reduced to a set of linear moments that can be estimated using either the generalized method of moments (GMM) directly or the two-pass estimation ("beta pricing") approach that evolved for linear pricing models prior to the explicit introduction of GMM estimators. As Cochrane (2001) and Jagannathan and Wang (2002) proved, these two approaches are mathematically identical as long as the GMM moment conditions are chosen appropriately.

Step #4: Construct the (standard) test statistics used to evaluate linear pricing models. The first step in this estimation scheme is a multivariate time-series regression of excess returns on the model factors:

$$\mathbf{R}^e_{t+1} = \mathbf{a}_0 + \mathbf{B}\mathbf{F}_{t+1} + \eta_{t+1}, \tag{40}$$

for $t = 1, \ldots, T$, where $\mathbf{F}_{t+1} \equiv (\mathbf{f}_{t+1} \otimes \mathbf{Z}_t)$, $\mathbf{B}$ is the $N \times KL$ matrix of regression coefficients (betas) and η_{t+1} is an N-vector of residual returns. The second step in the estimation is a cross-sectional regression consistent with the following unconditional moments for excess returns

$$E\left(\mathbf{R}^e_{t+1}\right) = \mathbf{B}\lambda + \alpha, \tag{41}$$

where λ is an LK-vector of constant risk premiums, and α is a vector of pricing errors. The test statistic for evaluating the joint significance of *all* pricing errors is

$$T\widehat{\alpha}'\widehat{\mathbf{\Sigma}}_{\alpha}^{-1}\widehat{\alpha} \overset{a}{\sim} \chi^2_N, \tag{42}$$

where $\overset{a}{\sim}$ means converges in distribution and $\widehat{\mathbf{\Sigma}}_{\alpha}$ is the covariance matrix for the pricing errors.[30] The significance of the risk premium estimates can be tested directly using the statistic

$$\widehat{\lambda}'\widehat{\mathbf{\Sigma}}_{\lambda}^{-1}\widehat{\lambda} \overset{a}{\sim} \chi^2_{KL}, \tag{43}$$

where $\widehat{\mathbf{\Sigma}}_{\lambda}$ is the covariance matrix for the risk premiums.

[30] Appendix A contains the complete expression for all covariance matrices and test statistics.

Individual pricing errors and individual risk premiums can be tested using the square root of the elements on the main diagonal of $\widehat{\mathbf{\Sigma}}_\alpha$ or $\widetilde{\mathbf{\Sigma}}_\lambda$, respectively. These models are also commonly evaluated heuristically by examining the size (and structure) of individual pricing errors and by computing average cross-sectional R-squared statistics.

B Covariance Matrices

The covariance matrix for α is defined as:

$$\widehat{\mathbf{\Sigma}}_\alpha = T^{-1}\left[\widehat{\mathbf{\Sigma}} - \widehat{\mathbf{B}}\left(\widehat{\mathbf{B}}'\widehat{\mathbf{\Sigma}}_\eta^{-1}\widehat{\mathbf{B}}\right)^{-1}\widehat{\mathbf{B}}'\right]\left(1+\widehat{\lambda}'\widehat{\mathbf{\Sigma}}_{\mathbf{F}}^{-1}\widehat{\lambda}\right), \tag{44}$$

where $\widehat{\mathbf{\Sigma}}_\eta$ is the sample estimate of the covariance matrix for η, $\widehat{\mathbf{\Sigma}}_{\mathbf{F}}$ is the sample estimate of the covariance matrix of the factors $\mathbf{F}$. The term in parentheses is a correction factor due to Shanken (1992) that accounts for the fact that the cross-sectional regressors are generated from the first-pass time series estimation.

If OLS is used in the cross-sectional regression, then there are two commonly used versions of the covariance matrix of the risk premiums. The first estimate is

$$\widehat{\mathbf{\Sigma}}_\lambda = T^{-1}\left[\left(\widehat{\mathbf{B}}'\widehat{\mathbf{B}}\right)^{-1}\widehat{\mathbf{B}}'\widehat{\mathbf{\Sigma}}_\eta\widehat{\mathbf{B}}\left(\widehat{\mathbf{B}}'\widehat{\mathbf{B}}\right)^{-1} + \widehat{\mathbf{\Sigma}}_{\mathbf{F}}\right], \tag{45}$$

where $\widehat{\mathbf{\Sigma}}_\eta$ is the estimated covariance matrix of the time series shocks and $\widehat{\mathbf{\Sigma}}_{\mathbf{F}}$ is the estimated factor covariance matrix. The second estimate uses the Shanken correction

$$\widetilde{\mathbf{\Sigma}}_\lambda = T^{-1}\left[\left(\widehat{\mathbf{B}}'\widehat{\mathbf{B}}\right)^{-1}\widehat{\mathbf{B}}'\widehat{\mathbf{\Sigma}}_\eta\widehat{\mathbf{B}}\left(\widehat{\mathbf{B}}'\widehat{\mathbf{B}}\right)^{-1}\left(1+\widehat{\lambda}'\widehat{\mathbf{\Sigma}}_{\mathbf{F}}^{-1}\widehat{\lambda}\right) + \widehat{\mathbf{\Sigma}}_{\mathbf{F}}\right]. \tag{46}$$

There are analogs to (45) and (46) when GLS is used in the cross-sectional regression.[31]

[31] The GLS estimators are

$$\widehat{\mathbf{\Sigma}}_\lambda = T^{-1}\left[\left(\widehat{\mathbf{B}}'\widehat{\mathbf{\Sigma}}_\eta^{-1}\widehat{\mathbf{B}}\right)^{-1} + \widehat{\mathbf{\Sigma}}_{\mathbf{F}}\right],$$

and

$$\widetilde{\mathbf{\Sigma}}_\lambda = T^{-1}\left[\left(\widehat{\mathbf{B}}'\widehat{\mathbf{\Sigma}}_\eta^{-1}\widehat{\mathbf{B}}\right)^{-1}\left(1+\widehat{\lambda}'\widehat{\mathbf{\Sigma}}_{\mathbf{F}}^{-1}\widehat{\lambda}\right) + \widehat{\mathbf{\Sigma}}_{\mathbf{F}}\right].$$

C The Investment Shock Proxy

Papanikolaou (2012) (equations (29) and (3) and the discussion in Section V) and Kogan and Papanikolaou (2012) (equation (26) and Section 3.1) establish the correspondence in the model between the embodied investment-specific technology shock and the return spread between investment and consumption sector portfolios. Essentially, this portfolio serves as the projection of the non-marketed factor onto the space of marketed returns.

In order to estimate a conditional model using a conventional cross-section of test assets, we need a longer time series of the spread returns than those used in Papanikolaou (2012) and Kogan and Papanikolaou (2012).[32] We use the 10 industry return series available from Ken French's website (at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html). The investment sector return is constructed as the equally weighted average of the value-weighted returns to the manufacturing, energy, hi-tech, and telecom industries. The consumption sector returns are constructed as the equally weighted average of the value-weighted returns to the consumer durables, consumer nondurables, and shops industries.[33]

A comparison of the constructed return differences, at the annual frequency, between our spread returns and the spread returns used in Papanikolaou (2012) and Kogan and Papanikolaou (2012) over the period from 1964 to 2008 (matching the earlier papers) is shown below.

[32] See Ferson and Foerster (1994).

[33] See French's website for the mapping between the industry definitions and SIC codes.

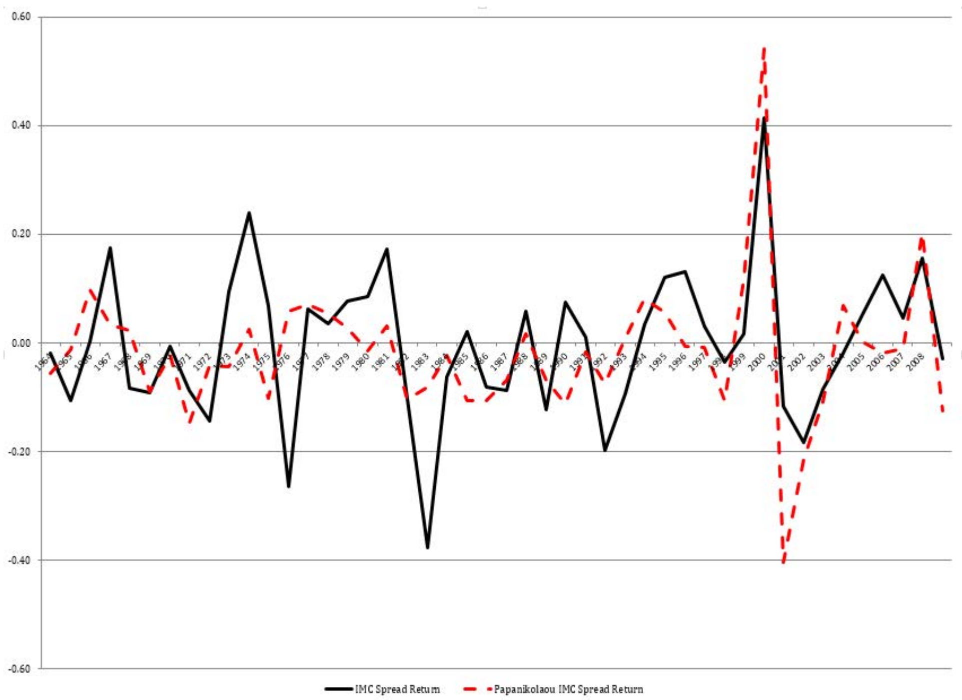

Figure C-1: The Two Time Series for the Differences in the Returns to Investment versus Consumption Sector Firms.

The series appear to move together, and they have an unconditional correlation coefficient of 0.59. Monthly continuously compounded returns are converted to quarterly by multiplying one plus the monthly returns within the quarter.

References

1. Bansal, R., D. A. Hsieh, and S. Viswanathan, 1993, "A New Approach to International Arbitrage Pricing," *Journal of Finance, 48*, pages 1719-1747.

2. Bansal, R., and A. Yaron, 2004, "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *Journal of Finance, 59*, pages 1481-1509.

3. Black, F., E. Derman, and W. Toy, 1990, "A One-Factor Model of Interest Rates and Its Application to Treasury Bond Options," *Financial Analysts Journal, 46*, pages 33-39.

4. Black. F., and P. Karasinski, 1991, "Bond and Option Pricing When Short Rates are Lognormal," *Financial Analysts Journal, 47*, pages 52-59.

5. Brandt. M. W., and Q. Kang, 2004, "On the Relationship Between the Conditional Mean and Volatility of Stock Returns: A Latent VAR Approach," *Journal of Financial Economics, 72*, pages 217-257.

6. Breeden, D. T., 1979, "An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities," *Journal of Financial Economics, 7*, pages 265-296.

7. Breeden, D. T., and M. R. Gibbons, and R. H. Litzenberger, 1989, "Empirical Tests of the Consumption-Oriented CAPM," *Journal of Finance, 44*, pages 231-262.

8. Brennan, M. J., A. W. Wang, and Y. Xia, 2004, "Estimation and Test of a Simple Model of Intertemporal Capital Asset Pricing," *Journal of Finance, 59*, pages 1743-1775.

9. Campbell, J. Y., 1987, "Stock Returns and the Term Structure," *Journal of Financial Economics, 7*, pages 265-296.

10. Campbell, J. Y., and J. H. Cochrane, 2000, "Explaining the Poor Performance of Consumption-Based Asset Pricing Models," *Journal of Finance, 55*, pages 2863-2878.

11. Chapman, D. A., 1997, "Approximating the Asset Pricing Kernel," *Journal of Finance, 52*, pages 1383-1410.

12. Chen, N., R. Roll, and S. A. Ross, 1986, "Economic Forces and the Stock Market," *Journal of Business, 59*, pages 383-403.

13. Cochrane, J. H., 1996, "A Cross-Sectional Test of an Investment-Based Asset Pricing Model," *Journal of Political Economy, 104*, pages 572-621.

14. Cochrane, J. H., 2001, "A Resurrection of the Stochastic Discount Factor/GMM Methodology," National Bureau of Economic Research, Working Paper No. 8533.

15. Cochrane, J. H., 2005, *Asset Pricing, Revised Edition.* (Princeton NJ: Princeton University Press.)

16. Dai, Q., K. J. Singleton, and W. Yang, 2007, "Regime Shifts in a Dynamic Term Structure Model of U.S. Treasury Bond Yields," *Review of Financial Studies, 20*, pages 1669-1706.

17. Engle, R., 2002, "Dynamic Conditional Correlation – A Simple Class of Multivariate GARCH Models," *Journal of Business and Economic Statistics, 20*, pages 339-350.

18. Fama, E. F., and K. R. French, 1993, "Common Risk Factors in the Returns on Stocks and Bonds," *Journal of Financial Economics, 33*, pages 3-56.

19. Ferson, W. E., and S. R. Foerster, 1994, "Finite Sample Properties of the Generalized Methods of Moments Tests of Conditional Asset Pricing Models," *Journal of Financial Economics, 36*, pages 29-56.

20. Ferson, W. E., and C. R. Harvey, 1991, "The Variation of Economic Risk Premiums," *Journal of Political Economy, 99*, pages 385-415.

21. Ferson, W. E., and C. R. Harvey, 1999, "Conditioning Variables and the Cross Section of Expected Stock Returns," *Journal of Finance 54*, pages 1325-1360.

22. Ferson, W. E., S. Sarkissian, and T. T. Simin, 2003, "Spurious Regressions in Financial Economics?" *Journal of Finance, 58*, pages 1393-1413.

23. Ghysels, E., 1998, "On Stable Factor Structures in the Pricing of Risk: Do Time-Varying Betas Help or Hurt?" *Journal of Finance, 53*, pages 549-573.

24. Gibbons, M. R., and W. E. Ferson, 1985, "Testing Asset Pricing Models with Changing Expectations and an Unobservable Market Portfolio," *Journal of Financial Economics, 14*, pages 216-236.

25. Gibbons, M. R., S. A. Ross, and J. Shanken, 1989, "A Test of the Efficiency of a Given Portfolio," *Econometrica, 57*, pages 1121-1152.

26. Graybill, F. A., 1984, *Matrices with Applications in Statistics, 2nd Ed.* Belmont, CA: Wadsworth International Group.

27. Hansen, L. P., and S. F. Richard, 1987, "The Role of Conditioning Information in Deducing Testable Restrictions Implied by Dynamic Asset Pricing Models," *Econometrica, 55*, pages 587-613.

28. Hansen, L. P., and K. J. Singleton, 1983, "Stochastic Consumption, Risk Aversion, and the Temporal Behavior of Asset Prices," *Journal of Political Economy, 91*, pages 249-268.

29. Harvey, C. R., 1989, "Time-Varying Conditional Covariances in Tests of Asset Pricing Models," *Journal of Financial Economics, 24*, pages 289-317.

30. Jagannathan, R., and Z. Wang, 1996, "The Conditional CAPM and the Cross-Section of Stock Returns," *Journal of Finance, 51*, pages 3-53.

31. Jagannathan, R., and Z. Wang, 2002, "Empirical Evaluation of Asset Pricing Models: A Comparison of the SDF and Beta Methods," *Journal of Finance, 57*, pages 2337-2367.

32. Judd, K. L., 1998, *Numerical Methods in Economics.* MIT Press.

33. Kan, R., C. Robotti, and J. Shanken, 2012, "Pricing Model Performance and the Two-Pass Cross-Sectional Regression Methodology," forthcoming in the *Journal of Finance.*

34. Kogan, L., and D. Papanikolaou, 2012, "Growth Opportunities, Technology Shocks, and Asset Prices," forthcoming in the *Journal of Finance.*

35. Kogan, L., D. Papanikolaou, and N. Stoffman, 2012, "Technological Innovation: Winners and Losers," manuscript, MIT, Northwestern, and Indiana University.

36. Lettau, M., and S. Ludvigson, 2001, "Resurrecting the (C)CAPM: A Cross-Sectional Test When Risk Premia Are Time-Varying," *Journal of Political Economy, 109*, pages 1238-1287.

37. Lewellen, J., and S. Nagel, 2006, "The Conditional CAPM Does Not Explain Asset Pricing Anomalies," *Journal of Financial Economics, 82*, pages 289-314.

38. Lustig, H., and S. van Nieuwerbergh, 2005, "Housing Collateral, Consumption Insurance, and Risk Premia: An Empirical Perspective," *Journal of Finance, 60*, pages 1167-1219.

39. Merton, R. C., 1973, "An Intertemporal Capital Asset Pricing Model," *Econometrica, 41*, 867-887.

40. Papanikolaou, D., 2011, "Investment Shocks and Asset Prices," *Journal of Political Economy, 119*, pages 639-685.

41. Roussanov, N. I., 2012, "Composition of Wealth, Conditioning Information, and the Cross-Section of Stock Returns," forthcoming in the *Journal of Financial Economics.*

42. Shanken, J., 1990, "Intertemporal Asset Pricing," *Journal of Econometrics, 45*, pages 99-120.

43. Shanken, J., 1992, "On the Estimation of Beta Pricing Models," *Review of Financial Studies, 5*, pages 1-34.

44. Santos, T. and P. Veronesi, 2006, "Labor Income and Predictable Stock Returns," *Review of Financial Studies, 19*, pages 1-44.

45. Yogo, M., 2006, "A Consumption-Based Explanation of Expected Stock Returns," *Journal of Finance, 61*, pages 539-580.