NBER WORKING PAPER SERIES

LINEAR APPROXIMATIONS AND TESTS
OF CONDITIONAL PRICING MODELS

Michael W. Brandt
David A. Chapman

Working Paper 12513
http://www.nber.org/papers/w12513

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2006

Linear Approximations and Tests of Conditional Pricing Models
Michael W. Brandt and David A. Chapman
NBER Working Paper No. 12513
September 2006
JEL No. G0, G10, G12, G14

ABSTRACT

We construct a simple reduced-form example of a conditional pricing model with modest intrinsic nonlinearity. The theoretical magnitude of the pricing errors (alphas) induced by the application of standard linear conditioning are derived as a direct consequence of an omitted variables bias. When the model is calibrated to either characteristics sorted or industry portfolios, we find that the alphas generated by approximation-induced specification error are economically large. A Monte Carlo analysis shows that finite-sample alphas are even larger. It also shows that the power to detect omitted nonlinear factors through tests based on estimated risk premiums can sometimes be quite low, even when the effect of misspecification on alphas is large.

Michael W. Brandt
Fuqua School of Business
Duke University
Box 90120
One Towerview Drive
Durham, NC 27708
and NBER
mbrandt@duke.edu

David A. Chapman
Finance Department
Boston College
140 Commonwealth Avenue
Fulton Hall 330
Chestnut Hill, MA 02467
david.chapman@bc.edu

1   Introduction

The distinction between a model that explains asset returns on average and one that explains returns conditional on the set of information used by market participants is at the heart of asset pricing theory. It is also of considerable practical importance. For example, Ferson and Harvey (1999) argue that the commonly used three-factor model first examined in Fama and French (1993) does not explain the cross-section of size and book-to-market sorted portfolios in a setting that conditions on common macroeconomic predictor variables. As another example, Campbell and Cochrane (2000) suggest that the distinction between a conditional and an unconditional pricing model is consistent with the superior performance of portfolio-based pricing models relative to theoretically more appealing consumption-based models.

It is conceptually difficult to work with fully structural conditional pricing models because of theoretical arguments first stated in Hansen and Richard (1987). This fundamental criticism of empirical work on conditional pricing models notes that these models are inherently dependent on the dynamics of the unobservable information structure in the economy. As a result, they do not "condition down" to produce valid unconditional models. The standard approach to estimating conditional pricing models is explicitly empirical and ignores this issue entirely.[1] It works by explicitly allowing for conditional elements of the asset pricing model through a specification that is linear in some set of observable variables.

Clearly, the success of this approach will be model dependent – as the Hansen-Richard critique dictates. For example, if the true joint distribution of returns and the pricing kernel is lognormal and the conditioning variables are conditionally Gaussian (and observable), then the approximation would, in fact, be an exact representation of the true pricing kernel (up to a potential "proxy" problem). On the other hand, if asset markets are subject to parameter instability or structural breaks – both extreme forms of nonlinearity – then the linear approximation would clearly be problematic and unlikely to be considered by

---

[1]The argument, is made explicitly in Cochrane (1996, 2001). The roots of this approach are found in Campbell (1987), Gibbons and Ferson (1985), Harvey (1989), and Shanken (1990).

a reasonable researcher. What about an intermediate case? How well does the standard approximation work in a model where the nonlinearity is modest and might, therefore, be difficult to detect?

We construct a simple and tractable conditional model for the dynamics of a broad cross-section of stock returns that builds on the intertemporal CAPM approach introduced in Brennan, Wang, and Xia (2004). The pricing kernel is driven by three factors: the innovations to the two (latent) state variables, the log short rate and the log of the maximum Sharpe ratio, and the innovations to the market portfolio. The model is intended as a parsimonious reduced-form example of an economy with modest nonlinearity. Structural specifications that might introduce nonlinearities are the preference-based approach of Dittmar (2002) or the heterogeneous agent model of Lochstoer (2006). An alternative reduced-form motivation for modest nonlinearity can also be found in the recent generalizations of the affine pricing models commonly used in the term structure literature.

The model is conditional in the commonly accepted definition of that term: the pricing kernel parameters are functions of the state variables.[2] The single source of nonlinearity in the model is the risk premium on the log Sharpe ratio state variable (the only risk premium in the model). While the risk premium is nonlinear, it can be approximated very accurately by a second order polynomial in the log Sharpe ratio. This approximation is important because it allows us to retain a pricing structure that is linear in the model parameters, and it allows us to interpret the misspecification introduced by the standard linear approximation of the conditional model as an omitted variables bias. We take pains to emphasize that the nonlinearity imposed on the risk premium function is both modest and consistent with return data.

In order to examine the extent of this bias, we first derive a closed-form expression for

---

[2]This approach to conditional pricing models is conceptually distinct from the issue of how to use conditioning information to augment the space of asset returns used to test a pricing model. See, for example, Gallant, Hansen, and Tauchen (1990), Ferson and Siegel (2001), and Bekaert and Lui (2004). These conditioning methods may be used to test both conditional and unconditional pricing models.

the alphas that are generated by the use of the linear approximation, in large samples. The quantitative significance of these pricing errors is evaluated by calibrating the model to two different, commonly used, data sets of U.S. monthly excess returns: market capitalization (size) and book-to-market (B/M) sorted portfolios and industry portfolios, from January 1960 to December 2005.

Our first result is that alphas due to misspecification are large relative to average excess returns, and they vary cross-sectionally. This is true in both size and B/M sorted returns and in industry returns. For example, in the lowest B/M stocks in each size quintile, the alpha that can be attributed to ignoring nonlinear components in the risk premium alone averages six percent per year. These "misspecification alphas" decline monotonically with increases in B/M within each size quintile, but they still average (roughly) 2.4 percent per year in the middle B/M group within each size quintile. The largest third of the industry misspecification alphas average (roughly) 3.5 percent per year. These cross-sectional differences are examined by decomposing asymptotic alphas into various components that reflect both the bias in estimated betas that comes from omitted variables and the mismeasurement of factor risk premiums.

The fact that these pricing errors are large is important, but it does not necessarily provide an accurate measure of the *finite sample* pricing errors that might arise from this misspecification. In order to address this question, we conduct two Monte Carlo experiments in which we simulate $5,000$ independent sample paths from the model economy – calibrated to each of the sets of portfolio returns. Simulated excess returns are generated from the true (nonlinear) model, with true alphas set to zero. The model is then fit using a standard two-pass estimation methodology that uses ordinary least squares in the second stage cross-sectional regression.[3] The results of these calculations demonstrate that finite sample pricing errors are routinely between 25 percent and 75 percent of average excess returns in the characteristics-sorted calibration and 15 to 25 percent of average excess returns in the

---

[3]All of our findings are robust to using generalized least squares in the cross-sectional regression.

industry calibration. As important, because of noise in estimating the risk premiums, these finite-sample alphas look very different from their large-sample counterparts.

An important question raised by these finding is: If the pricing effects of misspecification are of first-order importance, should the nonlinear components of the true set of factors not be priced in the market? If so, then detecting the presence of these nonlinear components would be straightforward. The final analysis in the paper addresses this specific issue. We use the same Monte Carlo experiments to examine the finite sample power of a standard chi-squared test of the null hypothesis that the risk premiums on the higher order terms in the model are zero.

These experiments indicate that the power of these standard tests can vary substantially, depending on both the data used to test the model and the manner in which the test is implemented. If the tests are implemented without a correction for the fact that betas are estimated in the first-pass time series regression, the test is very powerful in the simulations calibrated to the characteristics-sorted returns, but the test has very low power to detect nonlinearity in the risk premiums in the industry data. This is another piece of evidence consistent with the notion that the set of test assets is of first-order importance in testing alternative pricing models.[4] When the covariance matrix of the estimated risk premiums is adjusted for measurement error in the betas, then the power of the standard test is poor in both data sets.[5]

In summary, we have presented a simple example, calibrated to two different sets of actual excess returns, that demonstrates that even modest intrinsic nonlinearity can induce significant biases into estimates of model-implied alphas. Furthermore, this nonlinearity can be difficult to detect in samples that are similar to actual data sets, although this finding is very sensitive to both the data used to test the model and to the form in which the

---

[4]See also Daniel and Titman (2005) and Phalippou (2005).

[5]These findings are robust to whether the second-pass cross-sectional regression is implemented using ordinary least squares or generalized least squares or whether covariances are estimated using simple estimators or Newey-West robust covariance matrix estimators.

test is implemented. This example highlights the potential difficulties from approaching the inherent problem of modelling the dynamics of conditioning information through the "standard" approach.

2 A Reduced-Form Conditional Economy

2.1 The Pricing Kernel

We follow the approach in Brennan, Wang, and Xia (2004) (hereafter, BWX) for implementing a version of the Intertemporal Capital Asset Pricing Model (ICAPM) from Merton (1973). Financial markets are complete, and there are no arbitrage opportunities. The economy contains a stochastic investment opportunity set that is described by the return on the market portfolio and a pair of state variables: the level of the one-period nominal risk-free rate, and the slope of the capital market line; i.e., the Sharpe ratio on the (conditionally) mean-variance efficient portfolio that prices all assets. 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 (gross) yield on a one-period, default-risk free bond from t to $t + 1$. The dynamics of this 'short rate' are exogenously specified as

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

where $\varepsilon_{1t} \overset{iid}{\sim} \mathcal{N}(0, \sigma_1^2)$. The short rate process (1) was introduced to the term structure literature in Black, Derman, and Toy (1990) and Black and Karasinski (1991).

The existence of a valid (unique) asset pricing kernel implies the existence of a (unique) portfolio that is conditionally mean-variance efficient and prices all assets.[6] The Sharpe ratio on this asset is denoted s, and its dynamics are implied by

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

---

[6]See Hansen and Richard (1987) for a proof of this statement.

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

Combined, (1) and (2) is a linear, first-order vector autoregression in the logs of the state variables. (2) implies that the Sharpe ratio on the conditionally mean-variance efficient portfolio can never be negative. There is no specific structural model underlying (2), 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_1. In the following section, we will estimate the parameters of (1) and (2).

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

$$M_{t+1} = \exp\left(-r_t^f - \frac{1}{2}\Lambda_t^2 - \Lambda_t \omega^{-1}\xi_{t+1}\right), \tag{3}$$

where Λ_t is a risk premium parameter (discussed in more detail below), $\xi_{t+1} \equiv \boldsymbol{\delta}'\boldsymbol{\varepsilon}_{t+1}$, $\boldsymbol{\varepsilon}_{t+1} = (\varepsilon_{1t+1}, \varepsilon_{2t+1}, \varepsilon_{3t+1})'$, $\boldsymbol{\delta}$ is a vector of constants, $\omega = (\boldsymbol{\delta}'\Sigma\boldsymbol{\delta})^{1/2}$, and $\Sigma = I_3$ is the covariance matrix of $\boldsymbol{\varepsilon}_{t+1}$.

The pricing kernel (3) is in the exponential-affine form that is common in the asset pricing literature, particularly in the continuous-time term structure area. The single innovation to the kernel is a weighted-average of the innovations to the market return and the two state variables defined in (1) and (2). This is precisely the specification in BWX. The fact that the shock, ξ, is normally distributed implies that M_t is conditionally lognormal.

There are two economically significant restrictions imposed on the pricing kernel by (3). First, we assume a single risk premium on a single composite shock. This is purely 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.

The final assumption that we make in specifying the model economy is that asset returns from t to $t+1$ and the pricing kernel from t to $t+1$ 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 in the latent factors. Specifically,

$$E_t\left(r_{i,t+1}\right) + E_t\left(m_{t+1}\right) + \frac{1}{2}\left[\text{var}_t\left(r_{i,t+1}\right) + \text{var}_t\left(m_{t+1}\right) + 2\text{cov}_t\left(r_{i,t+1}, m_{t+1}\right)\right] = 0, \quad (4)$$

for $i = 1, \ldots, N$ and where $m_{t+1} \equiv \ln M_{t+1}$ and $r_{i,t+1} \equiv \ln R_{i,t+1}$, where $R_{i,t+1}$ is the gross return to asset i. Equation (4) can be rewritten in terms of returns, functions of latent factors, and factor innovations:

$$E_t\left(r_{i,t+1}\right) + \frac{1}{2}\text{var}_t\left(r_{i,t+1}\right) = r_t^f + \Lambda_t \text{cov}\left(r_{i,t+1}, \omega^{-1}\xi_{t+1}\right), \quad (5)$$

for $i = 1, \ldots, N$. The covariance in equation (5) 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 (5) can be rewritten in a traditional single-factor (conditional) beta pricing form:

$$E_t\left(r_{i,t+1}\right) + \frac{1}{2}\text{var}_t\left(r_{i,t+1}\right) - r_t^f = \Lambda_t \beta_i, \quad (6)$$

where

$$\beta_i = \frac{\text{cov}\left(r_{i,t+1}, \xi_{t+1}\right)}{\omega^2}. \quad (7)$$

As a final point, it is important to note that the risk premium parameter in equation (6) is identical to the pricing kernel coefficient in equation (3), which justifies our use of the

same notation. Although this is generally not true, see Cochrane (2005) for a discussion, it follows in this simple economy because the factor that drives the pricing kernel is *iid* standard normal.

2.2 The Risk Premium

The maximum conditional Sharpe ratio on all marketed assets, in any economy with a valid pricing kernel, is equal to the ratio of the conditional volatility of the pricing kernel to its conditional mean:

$$s_t = \frac{\sigma_t\left(M_{t+1}\right)}{E_t\left(M_{t+1}\right)} = \sqrt{\exp\left(\Lambda_t^2\right) - 1}, \tag{8}$$

where the second equality follows from the (conditional) lognormality of the pricing kernel.[7] Rearranging equation (8) yields

$$\Lambda_t = \sqrt{\ln\left(1 + s_t^2\right)}. \tag{9}$$

In constructing equation (9), we rely on the assumption that there is a single risk premium for the composite shock, ξ.

Figure 1 shows that the risk premium is a convex function of the level of the log Sharpe ratio. In order to simplify future analysis, it is useful to approximate equation (9) by a polynomial expansion. In particular we consider the first- and second-order approximations:

$$
\begin{aligned}
\widetilde{\Lambda}_t\left(1\right) &= \psi_0 + \psi_1 \ln s_t \\
\widetilde{\Lambda}_t\left(2\right) &= \psi_0 + \psi_1 \ln s_t + \psi_2 P_2\left(\ln s_t\right),
\end{aligned}
\tag{10}
$$

where ψ_i, $i = \{0, 1, 2\}$, are parameters of the approximation that can be fit by a simple linear regression. $P_2\left(\ln s_t\right)$ is the component of $\left(\ln s_t\right)^2$ that is orthogonal to $\ln s_t$ and normalized to have zero mean and unit variance. In addition to the true Λ_t, Figure 1 also shows the

---

[7]The proof of (8) 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 lognormality of the pricing kernel in (3).

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 $\ln s_t$.

2.3 Summary

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* completely linear. Although there are two 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. It 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 sections, both asymptotically and in finite sample.

3 Approximating a Conditional Model

3.1 The Standard Approach

Since our goal is to understand the extent to which misspecification induced by the "standard approach" to dealing with conditional pricing models affects inference, we first need to define

it. The standard approach to estimating and testing a conditional pricing model has evolved over time in a series of papers, including Campbell (1987), Gibbons and Ferson (1985), Harvey (1989), Shanken (1990), and Cochrane (1996), culminating in the textbook treatment of Cochrane (2005). It is based on three steps:

Step #1: State the model in the form of the fundamental asset pricing equation:

$$E\left(M_{t+1}\mathbf{R}^e_{t+1} \mid \mathcal{F}_t\right) = \mathbf{0}, \tag{11}$$

where $\mathbf{R}^e_{t+1}$ is an N-vector of asset returns in excess of the return to a reference asset (typically a proxy for a short term risk-free bond) from t to $t+1$ that are used to test the model; M_{t+1} is the asset pricing kernel (or stochastic discount factor) that defines the pricing model; $\mathbf{0}$ is an N-vector of zeros, 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:

$$M_{t+1} = 1 + \mathbf{f}'_{t+1}\mathbf{b}_t, \tag{12}$$

as either an exact representation of the model being tested 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 $\mathcal{F}_t$.

Step #3: The essence of the linear conditioning approximation consists of two components. The information set, $\mathcal{F}_t$, can be reduced to a K-vector of observable variables $\mathbf{Z}_t$, and $\mathbf{b}_t$ can be written as

$$\mathbf{b}_t = \mathbf{B}\mathbf{Z}_t, \tag{13}$$

where $\mathbf{B}$ is a $K \times L$ matrix of constant coefficients.[8]

---

[8]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.

It is important to note that the linear assumption in equation (13) is distinct from the linear form of the pricing kernel. There is often no obvious reason to impose linearity in the true model state variables – much less linearity in the proxies Z – on the pricing kernel coefficients. Indeed, equation (8) is an example of a form of nonlinearity that is implied directly by the choice of state variables in the example economy. The use of the linearity assumption is a direct result of the econometrician's inability to specify a complete structure for the dynamics of investor's information set.

The implications of the fundamental asset pricing equation (11) 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 Jagannathan and Wang (2002) and Cochrane (2001) proved, these two approaches are mathematically identical as long as the GMM moment conditions are chosen appropriately.

We use the beta pricing approach for two (related) reasons. First, the excess returns on the sets of test assets that are commonly chosen in empirical work often exhibit high contemporaneous correlation. This can make some of the numerical calculations of the standard GMM approach numerically unstable for a large cross-section of assets. Second, this instability has driven many recent papers on conditional pricing models to use versions of the beta pricing approach; see, for example, Lettau and Ludvigson (2001) or Santos and Veronesi (2005).

The beta pricing approach is well-known, and it will only be presented here (briefly) in order to establish notation.[9] The first step in this estimation scheme is a multivariate time-series regression of excess returns on the model factors:

$$\mathbf{R}_{t+1}^e = \mathbf{a}_0 + \mathbf{B}\mathbf{F}_{t+1} + \boldsymbol{\eta}_{t+1}, \tag{14}$$

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

---

[9]Textbook treatments of this approach include Campbell, Lo, and MacKinlay (1997) and Cochrane (2005).

(betas) and $\boldsymbol{\eta}_{t+1}$ is an N-vector of shocks. Note that this is an expanded unconditional model that is being used to approximate the true conditional model.

The second step in this approach is a cross-sectional regression consistent with the following unconditional moment for excess returns

$$E\left(\mathbf{R}_{t+1}^e\right) = \mathbf{B}\boldsymbol{\lambda} + \boldsymbol{\alpha}, \tag{15}$$

where $\boldsymbol{\lambda}$ is an LK-vector of constant risk premiums, and $\boldsymbol{\alpha}$ is a vector of pricing errors. The pricing error α_i is commonly referred to as the alpha of portfolio i.

The standard errors for the individual pricing errors are calculated as the square root of the diagonal elements of the following matrix

$$\widehat{\Sigma}_{\boldsymbol{\alpha}} = T^{-1}\left[\widehat{\Sigma} - \widehat{\mathbf{B}}\left(\widehat{\mathbf{B}}'\widehat{\Sigma}_{\boldsymbol{\eta}}^{-1}\widehat{\mathbf{B}}\right)^{-1}\widehat{\mathbf{B}}'\right]\left(1 + \widehat{\boldsymbol{\lambda}}'\widehat{\Sigma}_{\mathbf{F}}^{-1}\widehat{\boldsymbol{\lambda}}\right), \tag{16}$$

where $\widehat{\Sigma}_{\boldsymbol{\eta}}$ is the sample estimate of the covariance matrix of the time-series shocks, $\boldsymbol{\eta}$, $\widehat{\Sigma}_{\mathbf{F}}$ is the sample estimate of the covariance matrix of the factors $\mathbf{F}$, and 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.

It is possible to examine the market-wide prices of risk to assess the statistical and economic magnitudes of estimated risk premiums.[10] 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

$$\widehat{\Sigma}_{\boldsymbol{\lambda}} = T^{-1}\left[\left(\widehat{\mathbf{B}}'\widehat{\mathbf{B}}\right)^{-1}\widehat{\mathbf{B}}'\widehat{\Sigma}_{\boldsymbol{\eta}}\widehat{\mathbf{B}}\left(\widehat{\mathbf{B}}'\widehat{\mathbf{B}}\right)^{-1} + \widehat{\Sigma}_{\mathbf{F}}\right], \tag{17}$$

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

$$\widetilde{\Sigma}_{\boldsymbol{\lambda}} = T^{-1}\left[\left(\widehat{\mathbf{B}}'\widehat{\mathbf{B}}\right)^{-1}\widehat{\mathbf{B}}'\widehat{\Sigma}_{\boldsymbol{\eta}}\widehat{\mathbf{B}}\left(\widehat{\mathbf{B}}'\widehat{\mathbf{B}}\right)^{-1}\left(1 + \widehat{\boldsymbol{\lambda}}'\widehat{\Sigma}_{\mathbf{F}}^{-1}\widehat{\boldsymbol{\lambda}}\right) + \widehat{\Sigma}_{\mathbf{F}}\right]. \tag{18}$$

---

[10]There is also a standard methodology for evaluating the statistical significance of the pricing errors directly.

There are analogs to (17) and (18) when GLS is used in the cross-sectional regression.[11]

The statistical significance of individual risk premium estimates can be examined using the square-root of the diagonal elements of $\widehat{\Sigma}_{\boldsymbol{\lambda}}$ (or $\widetilde{\Sigma}_{\boldsymbol{\lambda}}$). The statistical significance of groups of risk premium estimates can be evaluated using a standard quadratic form

$$\widehat{\boldsymbol{\lambda}}' \left[\mathbf{C} \widehat{\Sigma}_{\boldsymbol{\lambda}} \mathbf{C}' \right]^{-1} \widehat{\boldsymbol{\lambda}} \overset{a}{\sim} \chi_q^2, \tag{19}$$

or

$$\widehat{\boldsymbol{\lambda}}' \left[\mathbf{C} \widetilde{\Sigma}_{\boldsymbol{\lambda}} \mathbf{C}' \right]^{-1} \widehat{\boldsymbol{\lambda}} \overset{a}{\sim} \chi_q^2. \tag{20}$$

where q denotes the number of zero restrictions, and $\mathbf{C}$ is the $q \times KL$ matrix of zeros and ones that implements the zero restrictions under the null hypothesis; i.e.,

$$H_0 : \mathbf{C} \widehat{\boldsymbol{\lambda}} = \underset{q \times 1}{\mathbf{0}}. \tag{21}$$

The null that all coefficients are equal to zero can be tested by setting $\mathbf{C} = \mathbf{I}_{KL}$.

3.2 Applying this Approach to the Example Pricing Kernel

We assume that the econometrician examining data generated by the model in Section 2 has access to excess returns and historical values of the market portfolio and of the true state variables. The time series regression corresponding to equation (14) is:

$$R_{i,t+1}^e = a_i + b_{i,1} \xi_{t+1} + b_{i,2} (\ln s_t) + b_{i,3} \left[\xi_{t+1} \cdot (\ln s_t) \right] + \eta_{i,t+1}, \tag{22}$$

for $i = 1, \dots, N$. 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, $\xi_{t+1} \cdot (\ln s_t)$.

---

[11]The GLS estimators are

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

and

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

Given that $\xi_{t+1} \equiv \delta_1 \varepsilon_{1,t+1} + \delta_2 \varepsilon_{2,t+1} + \delta_3 \varepsilon_{3,t+1}$, (22) actually expands to

$$R_{i,t+1}^e = a_i + \delta_1 b_{i,1} \varepsilon_{1,t+1} + \delta_2 b_{i,1} \varepsilon_{2,t+1} + \delta_3 b_{i,1} \varepsilon_{3,t+1} + b_{i,2} \left(\ln s_t \right) \tag{23}$$

$$+ \delta_1 b_{i,3} \left[\varepsilon_{1,t+1} \cdot \left(\ln s_t \right) \right] + \delta_2 b_{i,3} \left[\varepsilon_{2,t+1} \cdot \left(\ln s_t \right) \right]$$

$$+ \delta_3 b_{i,3} \left[\varepsilon_{3,t+1} \cdot \left(\ln s_t \right) \right] + \eta_{i,t+1}.$$

Since the δ_i and the original b_i coefficients cannot be identified separately, it is notationally convenient to rewrite (23) as

$$R_{i,t+1}^e = a_i + d_{i,1} \varepsilon_{1,t+1} + d_{i,2} \varepsilon_{2,t+1} + d_{i,3} \varepsilon_{3,t+1} + d_{i,4} \left(\ln s_t \right) \tag{24}$$

$$+ d_{i,5} \left[\varepsilon_{1,t+1} \cdot \left(\ln s_t \right) \right] + d_{i,6} \left[\varepsilon_{2,t+1} \cdot \left(\ln s_t \right) \right] + d_{i,7} \left[\varepsilon_{3,t+1} \cdot \left(\ln s_t \right) \right] + \eta_{i,t+1}.$$

The second-pass cross-sectional regression is

$$\overline{R}_i^e = \sum_{j=1}^{7} \lambda_j \widehat{d}_{i,j} + \upsilon_i, \tag{25}$$

for $i = 1, \ldots, N$, where $\overline{R}_i^e$ is the (time-series) average excess return to portfolio i, $\widehat{d}_{i,j}$ are the fitted value of the coefficients from equation (24), and the fitted values of υ_i corresponds to the measured alpha for asset i.

In the context of this model, it is the linearity assumption in equation (13) that is incorrect. Given the effectiveness of the second-order linear approximation to the true risk premium, a more accurate specification is

$$\begin{aligned} R_{i,t+1}^e =\ & a_i + c_{i,1} \varepsilon_{1,t+1} + c_{i,2} \varepsilon_{2,t+1} + c_{i,3} \varepsilon_{3,t+1} + c_{i,4} \left(\ln s_t \right) + c_{i,5} P_2 \left(\ln s_t \right) \\ & + c_{i,6} \left[\varepsilon_{1,t+1} \cdot \left(\ln s_t \right) \right] + c_{i,7} \left[\varepsilon_{1,t+1} \cdot P_2 \left(\ln s_t \right) \right] \\ & + c_{i,8} \left[\varepsilon_{7,t+1} \cdot \left(\ln s_t \right) \right] + c_{i,9} \left[\varepsilon_{2,t+1} \cdot P_2 \left(\ln s_t \right) \right] \\ & + c_{i,10} \left[\varepsilon_{3,t+1} \cdot \left(\ln s_t \right) \right] + c_{i,11} \left[\varepsilon_{3,t+1} \cdot P_2 \left(\ln s_t \right) \right] + \eta_{i,t+1}, \end{aligned} \tag{26}$$

The difference between equations (24) and (26) is the omission of the relevant factors involving $P_2 \left(\ln s_t \right)$. There is a corresponding (expanded) cross-sectional regression that is analogous to (25).

14

This comparison makes the source of the misspecification clear. The questions that we address below are: (i) When the true pricing errors (alphas) are zero, what are the magnitudes of the true (or asymptotic) alphas introduced by misspecification of the conditional pricing relation? (ii) How do these true alphas compare with finite-sample estimates of alphas? Finally, (iii) does a test statistic of the form in equation (19) for omitted priced factors have any power to detect the omitted variables in a sample of realistic size of observations drawn from this model?

3.3 Alphas Generated by Misspecificaton

Given the simple structure of the pricing kernel (using the second order approximation to the true risk premium function), it is possible to compute a closed-form expression for the alphas generated by misspecification. These alphas are based on the true moments of the factors and excess returns. In a general setting (which we later specialize to the assumptions of Section 2), let

$$\mathbf{F}_{t+1} = \left(\mathbf{F}'_{1,t+1}, \mathbf{F}'_{2,t+1}\right)', \tag{27}$$

where $\mathbf{F}_1$ is the L_1-vector of factors in the conventional approximation and $\mathbf{F}_2$ is the L_2-vector of higher-order terms omitted by the approximation, where $L = L_1 + L_2$. The covariance matrix of the factors is

$$\Sigma_{\mathbf{F}} = \begin{bmatrix} \Sigma_{11} & \vdots & \Sigma_{12} \\ \cdots & \vdots & \cdots \\ \Sigma_{21} & \vdots & \Sigma_{22} \end{bmatrix}, \tag{28}$$

where the partition is consistent with the partition of the factors. Σ_{11} denotes the covariance matrix of the included factors. Σ_{22} is the covariance matrix of the excluded factors, and $\Sigma_{12} = \Sigma'_{21}$ is the covariance matrix between the two sets of factors. The inverse of the factor covariance matrix is

$$\Sigma_{\mathbf{F}}^{-1} = \begin{bmatrix} \left(\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right)^{-1} & \vdots & -\Sigma_{11}\Sigma_{12}\left(\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\right)^{-1} \\ \cdots\cdots\cdots\cdots\cdots\cdots & \vdots & \cdots\cdots\cdots\cdots\cdots\cdots \\ -\Sigma_{22}\Sigma_{21}\left(\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right)^{-1} & \vdots & \left(\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\right)^{-1} \end{bmatrix}, \tag{29}$$

which follows from Theorem 8.2.1 in Graybill (1984).

The true betas in the correct model are

$$\mathbf{B} = \Sigma_{\mathbf{F}}^{-1}\Sigma_{\mathbf{FR}^e}, \tag{30}$$

where $\Sigma_{\mathbf{FR}^e}$ is the $L \times N$ covariance matrix of the F factors with the N returns in $\mathbf{R}^e$. Using equation (29) (and partitioning $\Sigma_{\mathbf{FR}^e}$ conformably as $\left(\Sigma'_{\mathbf{F}_1\mathbf{R}^e}, \Sigma'_{\mathbf{F}_2\mathbf{R}^e}\right)'$), the betas can be written as

$$\mathbf{B} = \left[\begin{array}{c} \left(\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right)^{-1}\Sigma_{\mathbf{F}_1\mathbf{R}^e} - \Sigma_{11}\Sigma_{12}\left(\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\right)^{-1}\Sigma_{\mathbf{F}_2\mathbf{R}^e} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ -\Sigma_{22}\Sigma_{21}\left(\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right)^{-1}\Sigma_{\mathbf{F}_1\mathbf{R}^e} + \left(\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\right)^{-1}\Sigma_{\mathbf{F}_2\mathbf{R}^e} \end{array} \right]. \tag{31}$$

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

$$E\left(\mathbf{R}^{e\prime}\right) = \left[\begin{array}{ccc} \boldsymbol{\lambda}'_1 & \vdots & \boldsymbol{\lambda}'_2 \end{array} \right]\mathbf{B}, \tag{32}$$

or

$$\begin{aligned} E\left(\mathbf{R}^{e\prime}\right) &= \boldsymbol{\lambda}'_1\left(\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right)^{-1}\Sigma_{\mathbf{F}_1\mathbf{R}^e} \\ &\quad - \boldsymbol{\lambda}'_1\Sigma_{11}\Sigma_{12}\left(\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\right)^{-1}\Sigma_{\mathbf{F}_2\mathbf{R}^e} \\ &\quad - \boldsymbol{\lambda}'_2\Sigma_{22}\Sigma_{21}\left(\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right)^{-1}\Sigma_{\mathbf{F}_1\mathbf{R}^e} \\ &\quad + \boldsymbol{\lambda}'_2\left(\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\right)^{-1}\Sigma_{\mathbf{F}_2\mathbf{R}^e}. \end{aligned} \tag{33}$$

In equation (32), expected returns are expressed exactly as a function of the factors that are included in the conventional approximation, the factors that are excluded from the conventional approximation, and the covariance of these two components.

In the standard approximation, expected returns are written as

$$E\left(\mathbf{R}^{e\prime}\right) = \boldsymbol{\alpha}' + \widetilde{\boldsymbol{\lambda}}'_1\Sigma_{11}^{-1}\Sigma_{\mathbf{F}_1\mathbf{R}^e}, \tag{34}$$

where $\widetilde{\boldsymbol{\lambda}}_1$ is the estimated risk premiums on the included factors only. Combining equations

16

(33) and (34), the pricing errors generated by the approximation are

$$
\begin{aligned}
\boldsymbol{\alpha}' \; = \; & \underbrace{\boldsymbol{\lambda}_1' \left(\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} \right)^{-1} \Sigma_{\mathbf{F}_1\mathbf{R}^e}}_{(a)} \\
& \underbrace{-\boldsymbol{\lambda}_1'\Sigma_{11}\Sigma_{12} \left(\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} \right)^{-1} \Sigma_{\mathbf{F}_2\mathbf{R}^e}}_{(b)} \\
& \underbrace{-\boldsymbol{\lambda}_2'\Sigma_{22}\Sigma_{21} \left(\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} \right)^{-1} \Sigma_{\mathbf{F}_1\mathbf{R}^e}}_{(c)} \\
& \underbrace{+\boldsymbol{\lambda}_2' \left(\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} \right)^{-1} \Sigma_{\mathbf{F}_2\mathbf{R}^e}}_{(d)} \\
& \underbrace{-\widetilde{\boldsymbol{\lambda}}_1'\Sigma_{11}^{-1}\Sigma_{\mathbf{F}_1\mathbf{R}^e}}_{(e)}
\end{aligned}
\tag{35}
$$

where the labels (a) through (e) will prove useful below. This expression can only be simplified by placing additional structure on the covariance between $\mathbf{F}_1$ and $\mathbf{F}_2$.

What would α look like when the omitted factors are uncorrelated with the included factors; i.e., if $\Sigma_{12} = \Sigma_{21}' = 0$? In this case, (35) simplifies considerably

$$
\boldsymbol{\alpha} = \Sigma_{11}^{-1}\Sigma_{\mathbf{F}_1\mathbf{R}^e} \left(\boldsymbol{\lambda}_1 - \widetilde{\boldsymbol{\lambda}}_1 \right) + \Sigma_{22}^{-1}\Sigma_{\mathbf{F}_2\mathbf{R}^e}\boldsymbol{\lambda}_2.
\tag{36}
$$

The alpha measured in the linear approximation has two components. The first component reflects measurement error in the estimated risk premium on the included factors that arises if B_1 and B_2 are not orthogonal, and the second component equals the portion of expected returns that can be attributed to the omitted factors. The additional terms in (35) measure the impact of omitting correlated factors on the estimates of the betas of the included factors.

4 Two Calibrated Examples

4.1 Calibrating the State Variable Dynamics

The first step in calibrating the model is to construct proxies for the state variables $\ln r^f$ and $\ln s$. These state variables will be used with alternative sets of cross-sectional asset returns in examining the model's implications. We use the yield on a one-month Treasury bill, from

17

CRSP, as a proxy for the short rate r^f. The log short rate is then $\ln r^f$. The sample period is from January 1960 to December 2005. 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. For example, our results are unchanged if we use a short rate measure extracted from a multifactor model fit to the entire term structure of Treasury yields.[12]

Identifying the log Sharpe ratio is more difficult. The model in Section 2.1 specifies the state variable as the log Sharpe ratio on 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.

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 function based on a nonlinear extension of the Kalman filter. Smoothed estimates of the conditional moments are then constructed based on the full sample of return data.[13] Our estimate of the log Sharpe ratio is constructed from these smoothed estimates. This filtering procedure also recovers a (full sample) estimate of the shocks to the market portfolio, ε_3, that can be used in calibrating the pricing kernel.[14] A plot of the log Sharpe ratio is shown in the bottom panel of Figure 2.

Given the time series of the state variables, we fit equations (2) and (1) to recover estimates of the parameters that describe these dynamics. Since this system is linear in logs, all of the parameters are estimated by ordinary least squares (OLS) with robust

---

[12]Specifically, Brandt and Chapman (2005) fit an essentially affine $\mathbb{A}_1(3)$ model of the Treasury term structure – using the taxonomy for dynamic models introduced in Dai and Singleton (2000) and Duffee (2002) – to a standard data set of government bonds using a simulated moment estimator. The short rate implied by this model produces calibration results that are virtually identical to those generated by the 1-month Treasury yield.

[13]Brandt and Kang (2005) use data from January 1946 to December 1998.

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

standard errors used for inference. Table 1 shows these results. The log Sharpe ratio is significantly positively related to its own lagged value and 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. While the log short rate is persistent, it does not appear to contain a unit root in the autoregressive representation of the series over this sample period.[15]

The fitted residuals from equations (1) and (2) 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).[16] These 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 $(\varepsilon_1, \varepsilon_2)'$).

The pricing kernel in equation (3) is completely specified when Λ and δ, the parameters that weight the individual shocks to define the composite shock, are chosen. These parameter values depend on the set of test asset returns, and their selection is described in the next two subsections.

4.2 Calibrating Returns to Characteristics Sorted Portfolios

Our first calibration of the pricing kernel uses the returns to the 25 portfolios formed by Fama and French (1993) on the basis of market capitalization ("size") and book-to-market ratio ("B/M"). The monthly portfolio returns are value-weighted and continuously compounded

---

[15]The *t*-test reported in Table 2 for the null of $\beta_1 = 1$ is only suggestive of the absence of a unit root. There is a large literature on the failings of standard asymptotic theory in this estimation setting.

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

in excess of the continuously compounded yield on a one-month Treasury bill. The sample period is January 1960 to December 2005 (for a total of $T = 552$ monthly observations).

The first four sample moments of excess returns are reported in Table 2. Holding size constant, the average excess returns increase from low B/M to high B/M portfolios. In our sample period, there is not a consistent size effect for a given B/M rank; i.e., holding B/M rank constant, returns do not monotonically decrease as capitalization increases. The volatility declines with size, and it also declines with the B/M ratio but only up to the fourth B/M quintile. The returns to all series are negatively skewed and exhibit excess kurtosis.

The calibrated values of the (asymptotic) alphas from the model in Section 2, fit to the characteristics-sorted returns are constructed by computing the required inputs: (i) the factor covariance matrices, Σ_{11}, Σ_{22}, and Σ_{12}; (ii) the covariance matrices of the factors with the excess returns, $\Sigma_{\mathbf{F}_1\mathbf{R}^e}$ and $\Sigma_{\mathbf{F}_2\mathbf{R}^e}$; and (iii) the market risk premium parameters, $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$. In this setting, the included and omitted factors are:

$$
\begin{aligned}
\underset{(7\times1)}{\mathbf{F}_1} &= (\varepsilon_{1,t+1}, \varepsilon_{2,t+1}, \varepsilon_{3,t+1}, \ln s_t, \varepsilon_{1,t+1} \cdot \ln s_t, \varepsilon_{2,t+1} \cdot \ln s_t, \varepsilon_{3,t+1} \cdot \ln s_t)' \\
\underset{(4\times1)}{\mathbf{F}_2} &= (P_2(\ln s_t), \varepsilon_{1,t+1} \cdot P_2(\ln s_t), \varepsilon_{2,t+1} \cdot P_2(\ln s_t), \varepsilon_{3,t+1} \cdot P_2(\ln s_t))'
\end{aligned}
\tag{37}
$$

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. 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. Therefore, the parameters are estimated using simple OLS in the second-stage cross-sectional regression. In addition, we do not correct the alpha and risk premium standard errors for measurement errors in the first-stage beta estimates.[17]

---

[17]The value of the correction term in the characterstics-sorted returns is 10.2 (4.99 in the industry returns examined in the next sub-section). When the correction term is applied to the calculation of the covariance matrix, none of the alphas in Table 3 are statistically significantly different from zero. We believe that using the unadjusted covariance matrix estimates is conservative.

The estimates of alpha from the full model are shown in Table 3. The model does a good job of pricing six portfolios, where 'good' means that average alphas are less than 5 percent of average excess returns. There are another four portfolios where the average alphas are between 5 and 10 percent of average excess returns. There are also six portfolios where alphas exceed 20 percent of average excess returns, including the smallest size and B/M portfolio where the alpha is more than four times the level of the average excess returns. There are nine portfolios where the estimated alpha seems large, relative to the simple OLS standard errors (i.e., the t-statistic exceeds two).

In the following sections, when we examine the impact of misspecification of the conditioning information on model estimation, we remove the estimated alphas from Table 3 from the average excess returns in order to construct a set of returns that corresponds to our null hypothesis of zero alphas under the assumed pricing kernel. In other words, in the simulations, our model prices the set of test assets perfectly, by construction. Clearly, this is a simplification of the real world and the simple pricing kernel proposed in Section 2 does not price some of the test assets well, but neither does it dramatically underperform many existing pricing models. Furthermore, it is not the focus of our analysis to assume that this is *the* reduced-form pricing model. The parameters used in computing values for alpha described in equation (35) are shown in the table in the appendix (Table A-1).

The alphas generated by misspecifying the conditioning information (hereafter "misspecification alphas"), i.e., the alphas in equation (35), are shown in Table 4. We also compute relative alphas, defined as the ratio of misspecification alphas to the average excess returns net of the estimated alphas in Table 3. The misspecification alphas are predominantly positive, which means that omitting the higher-order terms in the pricing kernel generates estimates of expected returns that are too small relative to true expected excess returns. More importantly, these alphas are large and increasing in firm size, whether measured in isolation or relative to adjusted average excess returns. For example, the average of the absolute value of alpha ranges across the size quintiles from 0.21 percent per month (2.52

21

percent per year) to 0.34 percent per month (4.08 percent per year). On a relative basis, these pricing errors are more than 50 percent of adjusted excess returns for twelve of the twenty five portfolios, and they are more than 25 percent (in absolute value) of adjusted excess returns for 18 portfolios. The extent of these large pricing errors indicates that misspecification is not concentrated in the small, low B/M stocks alone.

Is there a simple explanation for the cross-sectional difference in the alphas due to the misspecification induced by the linear approximation? Unfortunately, the answer to this question appears to be no. Since the omitted (higher-order) factors are correlated with the linear terms (i.e., $\Sigma_{12} \neq 0$; see the appendix Table A-1), there are no convenient simplifications that make the components easy to interpret. The component (e) is the largest piece of the misspecification alpha (in absolute value). It increases (in absolute value) as size increases. For each size quintile, the absolute value of (e) decreases as B/M increases. Components (c) and (d) depend on the size of the risk premium on the higher order terms, λ_2. The sum of these components is large and positive. It decreases (slightly) as size increases, and within a size quintile, it decreases as B/M increases. (a) and (b) are the components that depend on the risk premium on the first-order terms, λ_1. The sum of these components are generally positive, increasing in size, and decreasing in B/M within a size quintile. The magnitude for the sum of these components is smaller than the sum of (c) and (d), suggesting that the higher order terms contribute more to the misspecification alpha.

4.3   Calibrating Returns to Industry Portfolios

The results in Section 4.2 are calibrated to only one set (of commonly used) test assets. How sensitive is the conclusion of the importance of approximation-based specification error to the choice of test assets? In order to examine this question, we also present results for a model calibration tuned to the monthly excess returns on 30 value-weighted industry portfolios.[18]

---

[18]For further details on the industry portfolio construction, see Ken French's web page at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/tw_5_ports.html.

Tables 5 through 7 are identical in structure to Tables 2 through 4.

The sample moments in Table 5 are generally similar to those for the characteristics-sorted portfolios in Table 2. Summary statistics describing the cross-section of average excess returns to the two sets of test assets are:

	Size and Book-to-Market	Industry
Minimum	−0.0677	0.1229
Maximum	0.9630	0.9401
Median	0.5760	0.3754
Mean	0.5462	0.3936
Std. Dev.	0.2588	0.1500

The characteristics-sorted portfolios have a larger spread in average excess returns, but this is primarily due to the very low realized returns on the smallest size and B/M sorted portfolio (with excess return of −0.0677 percent per month). There is still substantial spread in the realized industry returns.

The fitted alphas from the industry portfolios are shown in Table 6. They are about half positive and half negative. The point estimates of the alphas are roughly the same size as the alphas from the characteristics-sorted portfolios. If anything, they are slightly larger. The standard errors of the alpha estimates using industry returns are larger than their characteristics-sorted counterparts; so, there are fewer alphas that are statistically significant at conventional levels. The *relative* alphas for the industry portfolios are generally larger than those for the characteristics-sorted portfolios, although there is no outlier that corresponds to the smallest size and B/M sorted portfolio. On a relative basis, these pricing errors are more than 50 percent of excess returns for 2 of the thirty portfolios, which is a substantially smaller proportion than in the characteristics-sorted data. Fifteen of the portfolio have relative alphas in excess of 25 percent (in absolute value) of excess returns. The model parameters needed to calibrate the misspecification alphas for the industry portfolios are shown in Table A-1.

23

The misspecification alphas from the industry portfolio calibration are in Table 7. They do not have the extremes of the misspecification alphas from the characteristics-sorted data, but they are not small. For example, the 90-th percentile (75-th percentile) absolute value of alpha is 0.17 percent per month (0.12 percent per month) or 2.04 percent per year (1.44 percent per year). On a relative basis, 17 of the 30 (absolute) alphas are more than 50 percent of adjusted excess returns, and 25 of the (absolute) alphas are more than 25 percent of adjusted excess returns. These proportions exceed their counterparts in the characteristics-sorted data. The fact that there is such a large difference in the impact of approximation-induced specification error is an important finding for interpreting empirical tests of asset pricing models.[19]

In addition to the total misspecification alpha, Table 7 also reports the components of alpha. They are very different from their counterparts in Table 4, although (again) the intuition for these different components is unclear. (e) is no longer predominantly negative, and it is much smaller in absolute value than in Table 4. (d) remains predominantly positive, but it is also much smaller in absolute value than it is based on characteristics-sorted test assets. Finally, (b) switches from positive to negative, and (a) is much smaller in absolute value.

5   A Monte Carlo Study

There are two reasons for considering a Monte Carlo study at this point. First, the results in the previous section provided evidence about the magnitude of the *asymptotic* alphas that can be attributed to misspecification of the conditional model. Finite sample pricing errors may be quite different from these asymptotic values. This difference reflects, in part, the noise in realized returns (and its impact on the sampling distribution), but it also reflects the impact of measurement error – individually and through their interaction with the estimates

---

[19]Daniel and Titman (2005) and Phalippou (2005) critically examine the use of size- and book-to-market-sorted portfolios in asset pricing tests.

of the other components of (35).

The second reason to consider a finite sample analysis of the misspecification problem is the issue of detecting the presence of nonlinearity in the risk premiums or, equivalently (in this setting), the coefficients of the pricing kernel. Do common tests of the form of equation (19) in Section 3.1 have reasonable finite sample power? 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.

5.1   Simulating Returns from the Models

The general structure of the experiments that we consider below is:

Step #1: Generate $5,000$ independent simulated sample paths of an economy in which returns are generated according to a calibrated version of the model economy, using either the size and B/M sorted portfolio parameters or the industry portfolios parameters. 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 point estimates of the parameter values in Table 1 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. 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 diagonal elements of the covariance matrix of the shocks from equation (22).

25

Step #2: Each simulation is of length 550 months, and the state variables and realized excess returns are collected after a "burn-in" period of 4,950 months designed to reduce the influence of initial conditions.[20]

Step #3: For each simulated path:

Step #3a: we estimate a linearized version of the conditional model, as in equations (22) and (25). The model is estimated using a standard two-pass procedure with OLS in the second-stage cross-sectional regression.[21] The alphas (pricing errors) for each asset in each simulation are stored for later examination. The estimated risk premiums are also computed and stored.

Step #3b: We also estimate a version of an approximate model that includes second-order terms in the conditioning variables. We construct the test statistic in (19) and (20) 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 test against the specific alternative).

The first column of Tables 8 and 9 shows the average excess returns to the test portfolio (either characteristics or industry based). These are the cross-simulation average excess returns of the time series average excess returns along each simulated sample path. The standard deviation of excess returns, reported in the second column of each of these tables, is also the cross-simulation average of the time series standard deviation computed along each simulated path.

The simulated excess returns to the characteristics-sorted portfolios in Table 8 are reasonable. They are comparable in magnitude to the actual returns reported in Table 2, 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 simulated returns to the industry

---

[20]The unconditional means of the state variables are used as the initial conditions for each simulation.

[21]We also consider GLS estimation of the second stage cross-sectional regression. These results are qualitatively similar to the reported OLS results, and they are available on request.

portfolios, shown in the first two columns of Table 9, are also reasonable, producing average excess returns that are generally consistent with the actual industry returns reported in Table 5.

5.2 Examining Finite-Sample Alphas

The simulated alphas generated by the model in Section 2 calibrated to characteristic-sorted portfolios are shown in Table 8. They are large, in both absolute and relative terms. The alphas are largest (in absolute value) for the lowest B/M portfolio in four of the five size quintiles. The largest alphas are around 0.08 to 0.14 percent per month (0.96 to 1.68 percent per year). Relative to average expected returns, the largest alphas are between 50 percent and (nearly) 200 percent of average excess returns.

The cross-sectional standard deviation of average alphas is roughly constant across the different portfolios at 0.11 percent per month (1.32 percent per year). 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 far more variable, indicating that there are many sample paths where estimated alphas are very large relative to average excess returns. For the lowest B/M assets in each size quintile, the standard deviation of the distribution of relative standard errors ranges from 22 times average excess returns to more than 65 times average excess returns.

By construction, the true alphas are zero in the correctly specified model. However, the large-sample misspecification alphas are shown in the first column of Table 4. When the results in Table 8 are compared to these values, there is very little correspondence between the finite-sample and the large-sample misspecification-induced alphas. For example, in Table 4, the alphas in each size quintile are monotonically declining in the B/M level, with the difference between the lowest and highest B/M portfolios roughly on the order of 0.6 to 0.8 percent per month (7.2 to 9.6 percent per year). There is an attenuation bias in

27

the finite-sample misspecification alphas; i.e., they are much smaller in the cross-simulation averages than the large-sample misspecifcation alphas (roughly 0.2 percent per month in the smaller size quintiles). Furthermore, there is no monotonic relation between alphas and B/M level.

The source of this difference appears to be related to estimation of the risk premium parameters and their interaction with estimates of factor and error covariances. In separate simulations (not reported), the finite-sample alphas were computed using the true betas (for the terms included in the standard approximation), and they were very similar to the ones reported in Table 8.[22] This is important evidence about the effects of misspecification in finite-sample that differ from their large-sample counterparts.

The simulated alphas in Table 9, calibrated to the industry excess returns, are also not similar to their counterparts in the first column of Table 7, but this comparison is more difficult to see because there is not a clear pattern in the alphas in Table 7 that corresponds to the pattern in the characteristics-sorted portfolios. For example, most of the large-sample alphas are positive, whereas 16 of 25 of the averages of the simulated alphas are negative. The largest (in absolute value) of the alphas in Table 7 are roughly 0.4 percent per month (4.8 percent per year), whereas the largest of the averages of the simulated industry alphas are roughly 0.15 percent per month (or 1.8 percent per year).

The standard deviations of the cross-simulation distributions of the industry data are lower than the standard deviations in the characteristics-sorted data (0.08 percent per month versus 0.11 percent per month). The averages of the relative pricing errors are significant for many portfolios, with the largest average alphas on the order of 70 to 250 percent of average excess returns. The standard deviations of the relative pricing error distributions are, again, much larger than the standard deviations of the absolute pricing errors. In short, the two alternative calibrations generate consistent results: finite-sample alphas due to misspecification can be large and they seem to be different from large-sample alphas.

---

[22]These simulation results are available on request.

5.3   Testing for Nonlinear Terms Directly

The final question that we can address through the simulated data is how easy it is for the test statistic in equation (19) or (20) to detect the presence of higher-order nonlinear terms in an expanded conditional factor model. After all, if this test has power against a reasonable alternative (defined here by the model in Section 2), then misspecification can be easily detected directly from the estimated prices of risk of the model with nonlinear terms, and they need not, then, show up in the estimated pricing errors.

In light of the standard approximation approach, a reasonable specification for testing for the presence of higher order terms in the factor model implied by the pricing kernel is to include terms in the second power of the state variables. If we suppress cross-product terms, these means testing the null hypothesis

$$\mathbf{H}_0 : \lambda_{(\ln S_t)^2 \cdot \varepsilon_{1,t+1}} = \lambda_{(\ln S_t)^2 \cdot \varepsilon_{2,t+1}} = \lambda_{(\ln S_t)^2 \cdot \varepsilon_{3,t+1}} = \tag{38}$$
$$\lambda_{(\ln r_t)^2 \cdot \varepsilon_{1,t+1}} = \lambda_{(\ln r_t)^2 \cdot \varepsilon_{2,t+1}} = \lambda_{(\ln r_t)^2 \cdot \varepsilon_{3,t+1}} = 0;$$

i.e., the higher-order conditional terms are not priced in the excess returns. In the context of the full model, where $\mathbf{F} = (\mathbf{F}_1', \mathbf{F}_2')'$, the $\mathbf{C}$ matrix is simply

$$\mathbf{C} = \left[\begin{array}{ccc} \mathbf{0}_{6 \times 11} & \vdots & \mathbf{I}_6 \end{array} \right],$$

and the test is of the form of (19) or (20), which is asymptotically distributed as a chi-squared random variable with six degrees of freedom.

In order to assess the finite-sample power of (19) or (20) in this setting, we estimate an expanded conditional factor model with quadratic terms in the state variables along each of the sample paths used to generate the results in Tables 8 and 9. The frequency with which (19) or (20) rejects the false null, for different nominal sizes of the tests, are shown in Table 10.[23]

---

[23]The results based on using GLS in the cross-sectional regression are available on request. They are virtually identical to the results reported in Table 10 and discussed below.

In Panel A, we consider the version of the test that uses a covariance matrix that does not adjust for the estimation of the betas in the first-pass time series regression. There is a striking difference in the power of the test in the two different calibrations. The simulations based on the parameter values from the characteristics-sorted returns shows a test with a great deal of power, at all of the nominal size levels examined. For example, for a nominal size of 5 percent, the test rejected the false null 96.8 percent of the time (i.e., the test statistic exceeded the nominal critical value of 12.5916 on $4,840$ of the $5,000$ simulated paths). The results for the calibration tuned to the parameters based on industry returns is completely different. Here the test has virtually no power to detect the false null, and the actual rejection rates are substantially lower than the nominal size of the test. For example, when the 5 percent critical value is used in the industry return simulations, the false null is rejected on only 20 sample paths.

Panel B of Table 10 shows the effect of correcting for the stochastic betas using the adjustment derived in Shanken (1992). The two different parameterizations retain their relative ranking in the sense that the test appears to be more powerful in the version of the model calibrated to the characteristics-sorted return data. However, the Shanken adjustment results in a dramatic reduction in the power of the test in the characteristics-sorted simulations. Instead of rejection rates well in excess of 90 percent, the rejection of the false null goes from 20 percent (with a nominal size of 10 percent) to a little over 2 percent (when the nominal size is 1 percent). The power of the Shanken-corrected test is also uniformly lower in the industry calibration of the model, but since the power of the uncorrected test was so low in this setting, the difference are less dramatic.

These results follows because the Shanken correction term is very different across the two simulations. In the characteristics-sorted data, the average value (across the simulations) is 5.94, with a standard deviation of 2.91.[24] Therefore, applying the correction for stochastic betas generates a much larger covariance matrix for the test statistic. In the simulations

---

[24]These results are similar in GLS estimates, whether or not robust covariance matrices are used to estimate the components of the Shanken correction. These results are available on request.

calibrated to industry returns, the average value of the correction is only 1.74 with a standard deviation of 0.32.

6 Conclusions

How poorly does the standard approximation to a conditional model work in practice? This question is – by necessity – model specific. In this paper, we constructed a simple, reduced-form pricing model with modest intrinsic nonlinearity. We showed that, in this case, the linear approximation introduces an omitted variables bias that can be decomposed into terms that are due to biases introduced in estimated betas and components that primarily reflect omitted risk premium terms.

When these asymptotic alphas are calibrated to commonly used excess return data sets, we demonstrate that they can be very large. Finally, we show that nonlinearity in the risk premiums of a simple model can be very difficult to detect in finite sample even when they have an economically significant effect on pricing errors.

This example – precisely because it is not pathologically nonlinear – raises an important caveat about the use of the standard nonstructural approach to estimating conditional models. It suggests that structural modelling of conditional risk premium models may be important in pricing the cross-section of expected returns.
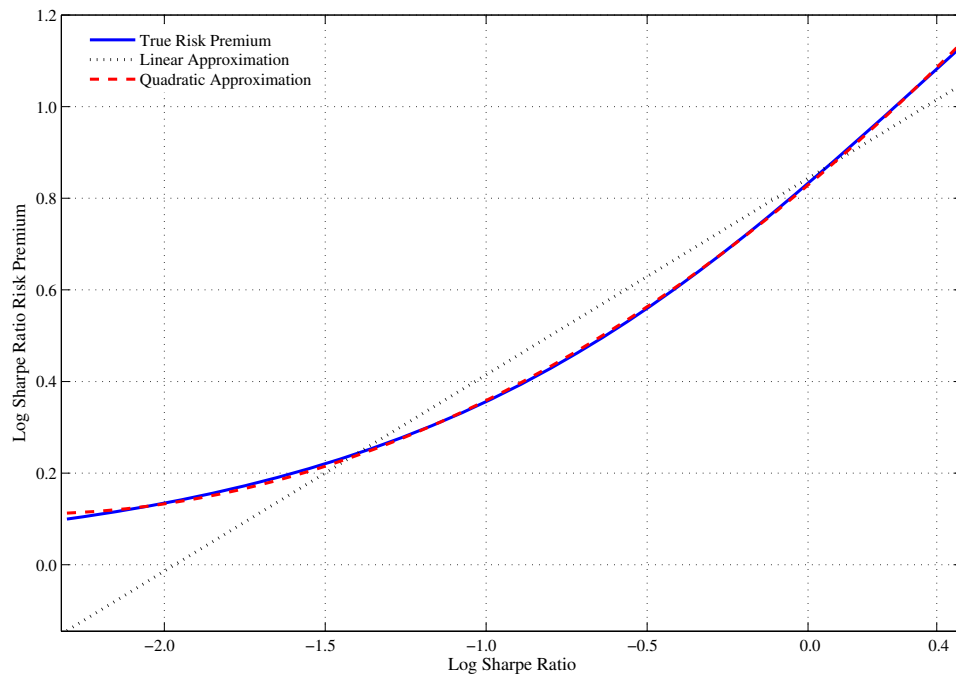
Figure 1: Log Sharpe Ratio Risk Premium and Approximations
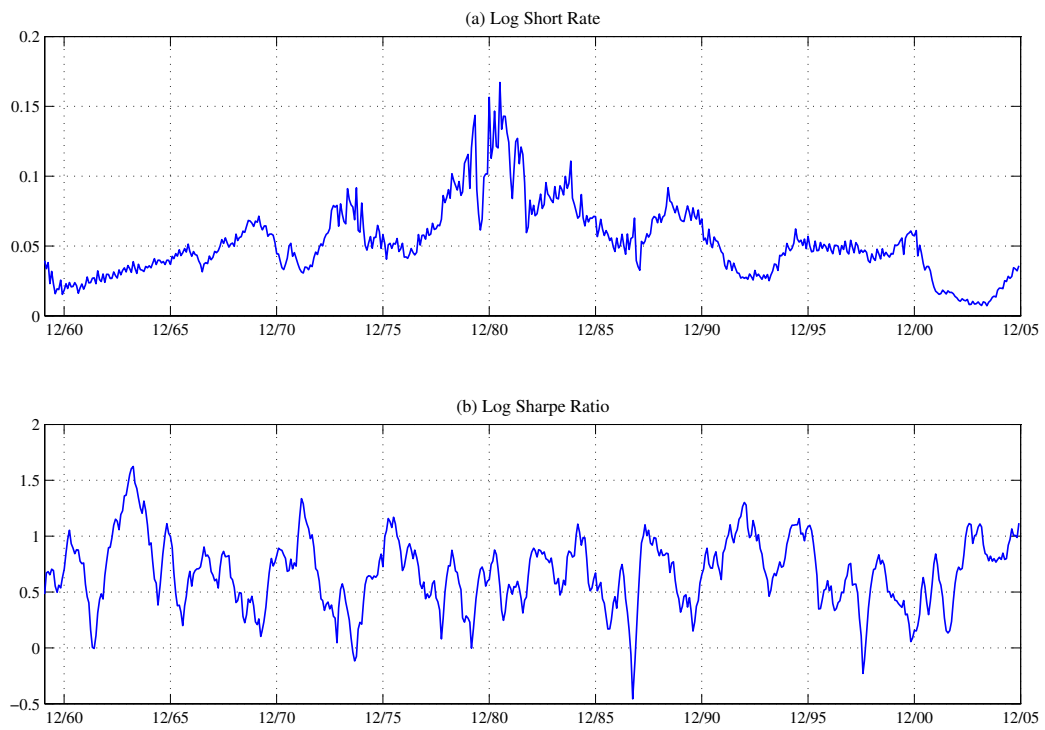
Figure 2: Pricing Kernel State Variables

Table 1: State Variable Parameter Estimates

Series	α_1	β_1		σ_1
	0.003	0.946		
ln (Short Rate)	(0.0007)	(0.014)		0.008
	[4.066]	[−3.934]		

Series	α_2	β_2	γ	σ_2
	0.069	0.930	−0.402	
ln (Sharpe Ratio)	(0.020)	(0.017)	(0.201)	0.106
	[3.445]	[−4.124]	[−2.003]	

The parameter estimates are for equations (1) and (2) in the text. The sample period is from January 1960 to December 2005 ($T = 552$ months). The parameters are estimated using ordinary least squares. Robust standard error estimates, based on the Newey-West estimator with 8 lags, are in parentheses. The brackets are t-tests for $\alpha_i = 0$, $i = 1, 2$, $\beta_i = 1$, for $i = 1, 2$, and $\gamma = 0$.

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

Portfolio	Mean	Std Dev	Skewness	Kurtosis
S1 BM1	−0.0677	8.2267	−0.4974	5.6265
S1 BM2	0.5187	6.9283	−0.4511	6.4128
S1 BM3	0.6494	5.9227	−0.5793	6.7574
S1 BM4	0.8891	5.5349	−0.5894	7.4351
S1 BM5	0.9630	5.8288	−0.5706	7.2118
S2 BM1	0.0929	7.4294	−0.6919	5.3717
S2 BM2	0.4631	6.0330	−0.8757	7.0726
S2 BM3	0.7183	5.3411	−0.9128	7.5597
S2 BM4	0.8309	5.1266	−0.7954	7.2121
S2 BM5	0.8652	5.7146	−0.7424	7.3289
S3 BM1	0.1755	6.7997	−0.6347	5.0977
S3 BM2	0.5767	5.4202	−0.9938	7.5412
S3 BM3	0.5760	4.9075	−0.9243	6.7773
S3 BM4	0.7181	4.7396	−0.5761	6.1058
S3 BM5	0.8225	5.3709	−0.7719	7.6272
S4 BM1	0.3335	5.9951	−0.4883	5.0226
S4 BM2	0.3873	5.1221	−0.9518	7.8582
S4 BM3	0.6285	4.8165	−0.7473	6.8756
S4 BM4	0.7448	4.6332	−0.3382	4.7288
S4 BM5	0.7041	5.2953	−0.5248	5.8456
S5 BM1	0.2824	4.7811	−0.4576	4.8760
S5 BM2	0.3820	4.5007	−0.6062	5.6129
S5 BM3	0.4460	4.2584	−0.5400	6.2279
S5 BM4	0.4712	4.1596	−0.1539	4.3159
S5 BM5	0.4833	4.7602	−0.3596	4.1915

Returns are continuously compounded in excess of the one-month Treasury bill yield, and they are measured over the period from January 1960 to December 2005. 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.).

Table 3: Fitted Alphas from the Full Model: Size and Book-to-Market Sorted Portfolios

Portfolio	α	StdErr(α)	t-statistic	Relative α
S1 BM1	-0.1969	0.0541	-3.6380	4.5903
S1 BM2	0.1822	0.0520	3.5032	0.3344
S1 BM3	0.0320	0.0422	0.7589	0.0481
S1 BM4	0.0211	0.0353	0.5981	0.0233
S1 BM5	0.0063	0.0329	0.1911	0.0064
S2 BM1	-0.0327	0.0383	-0.8534	-0.2838
S2 BM2	-0.0909	0.0450	-2.0182	-0.1863
S2 BM3	0.0931	0.0395	2.3581	0.1261
S2 BM4	0.0815	0.0436	1.8699	0.0955
S2 BM5	-0.1141	0.0467	-2.4436	-0.1283
S3 BM1	-0.0081	0.0474	-0.1718	-0.0404
S3 BM2	0.0526	0.0452	1.1634	0.0877
S3 BM3	-0.0494	0.0464	-1.0652	-0.0822
S3 BM4	0.0376	0.0435	0.8626	0.0507
S3 BM5	-0.0590	0.0500	-1.1796	-0.0697
S4 BM1	0.1240	0.0409	3.0336	0.3455
S4 BM2	-0.0719	0.0407	-1.7670	-0.1755
S4 BM3	0.2187	0.0484	4.5225	0.3340
S4 BM4	0.1196	0.0510	2.3449	0.1545
S4 BM5	-0.0274	0.0330	-0.8292	-0.0378
S5 BM1	-0.0400	0.0394	-1.0146	0.1272
S5 BM2	-0.0603	0.0445	-1.3564	-0.1473
S5 BM3	-0.0512	0.0549	-0.9311	-0.1101
S5 BM4	-0.0198	0.0395	-0.4998	-0.0398
S5 BM5	-0.1099	0.0439	-2.5027	-0.2145

Alphas are computed using the standard two-pass approach applied to the conditional model with a second-order approximation to the risk premium function; i.e., using equations (24) and (25) in the text. They are reported in percent per month. The data are from the period from January 1960 to December 2005 ($T = 552$ months). Standard errors for the alphas are constructed using equation (16) in the text. The t-statistic is the ratio of the estimate of alpha to its standard error. Relative alpha for a given portfolio is defined as the portfolio's alpha divided by the average excess return to the portfolio.

Table 4: Misspecification Alpha and Its Components: Size and Book-to-Market Sorted Portfolios

| Portfolio | α | α Components | | | | | Rel. α |
		(a)	(b)	(c)	(d)	(e)	
S1 BM1	0.485	−0.087	0.566	−0.332	2.712	−2.374	3.145
S1 BM2	0.245	−0.089	0.421	−0.267	2.150	−1.970	0.676
S1 BM3	0.079	−0.035	0.331	−0.245	1.794	−1.766	0.124
S1 BM4	0.156	0.015	0.294	−0.195	1.693	−1.652	0.177
S1 BM5	−0.106	−0.556	0.299	−0.285	1.714	−1.277	−0.109
S2 BM1	0.551	0.643	0.542	−0.224	2.513	−2.923	3.719
S2 BM2	0.206	0.337	0.388	−0.191	1.884	−2.212	0.356
S2 BM3	0.223	0.513	0.380	−0.208	1.683	−2.144	0.346
S2 BM4	0.039	0.332	0.339	−0.242	1.544	−1.933	0.051
S2 BM5	−0.306	0.273	0.344	−0.364	1.531	−2.090	−0.305
S3 BM1	0.519	1.042	0.525	−0.289	2.388	−3.147	2.447
S3 BM2	0.362	0.597	0.458	−0.212	1.845	−2.327	0.660
S3 BM3	0.251	0.480	0.415	−0.214	1.607	−2.036	0.385
S3 BM4	−0.112	0.577	0.333	−0.291	1.355	−2.085	−0.159
S3 BM5	−0.084	0.353	0.330	−0.263	1.483	−1.987	−0.093
S4 BM1	0.550	1.649	0.512	−0.279	2.144	−3.476	2.340
S4 BM2	0.448	1.170	0.473	−0.153	1.674	−2.716	0.929
S4 BM3	0.150	0.930	0.403	−0.236	1.459	−2.406	0.343
S4 BM4	−0.348	0.830	0.317	−0.404	1.256	−2.348	−0.532
S4 BM5	−0.117	0.717	0.369	−0.379	1.524	−2.348	−0.156
S5 BM1	0.491	1.881	0.425	−0.307	1.778	−3.285	1.387
S5 BM2	0.282	1.641	0.409	−0.287	1.487	−2.968	0.600
S5 BM3	0.379	1.894	0.389	−0.173	1.238	−2.969	0.734
S5 BM4	−0.168	1.510	0.261	−0.274	0.964	−2.629	−0.326
S5 BM5	−0.388	0.831	0.235	−0.301	0.980	−2.133	−0.624

This table shows the true values of the portfolio alphas generated by the misspecification induced by the standard linear approximation. Overall alpha (and its components) are quoted in percent per month. They are constructed using equation (35) in the text. The total alpha value is the sum of the different components:

$$
\begin{aligned}
(a) &= \boldsymbol{\lambda}_1' \left(\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} \right)^{-1} \Sigma_{\mathbf{F}_1\mathbf{R}^e} \\
(b) &= -\boldsymbol{\lambda}_1'\Sigma_{11}\Sigma_{12} \left(\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} \right)^{-1} \Sigma_{\mathbf{F}_2\mathbf{R}^e} \\
(c) &= -\boldsymbol{\lambda}_2'\Sigma_{22}\Sigma_{21} \left(\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} \right)^{-1} \Sigma_{\mathbf{F}_1\mathbf{R}^e} \\
(d) &= +\boldsymbol{\lambda}_2' \left(\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} \right)^{-1} \Sigma_{\mathbf{F}_2\mathbf{R}^e} \\
(e) &= -\widetilde{\boldsymbol{\lambda}}_1'\Sigma_{11}^{-1}\Sigma_{\mathbf{F}_1\mathbf{R}^e}
\end{aligned}
$$

Relative alpha is defined as true alpha divided by average excess returns net of the component of returns that the model cannot explain (α from Table 3).

Table 5: Summary Statistics for the Excess Returns to Industry Portfolios

Portfolio	Mean	Std Dev	Skewness	Kurtosis
Food Products	0.5424	4.4355	−0.2093	4.9323
Beer & Liquor	0.6335	5.3204	−0.5655	6.3490
Tobacco Products	0.9401	6.7306	−0.4698	5.5434
Recreation	0.3337	7.2852	−0.4889	4.9330
Printing & Publishing	0.3412	5.6538	−0.5369	5.1435
Consumer Goods	0.3787	4.8775	−0.4916	4.7537
Apparel	0.4062	6.5981	−0.6896	6.4706
Health Care, Med. Eq. & Pharm.	0.4889	5.0341	−0.2729	5.0589
Chemicals	0.2630	5.2686	−0.3838	6.3571
Textiles	0.2744	6.0472	−0.9398	7.1470
Construction & Materials	0.3762	5.4926	−0.7096	6.4672
Steel Works Etc.	0.1229	6.6510	−0.4293	5.8689
Fabricated Products & Machinery	0.3556	5.9600	−0.6335	5.5095
Electrical Equipment	0.3358	5.6077	−0.3278	4.6439
Automobiles & Trucks	0.1828	6.0170	−0.4796	5.3950
Aircraft, Ships, & Railroad Eq.	0.3888	6.4745	−0.6440	5.3635
Precious Metals & Mining	0.3943	7.0262	−0.2972	5.3116
Coal	0.5003	8.2698	0.0920	5.5349
Petroleum & Natural Gas	0.5216	5.1079	−0.0726	4.6175
Utilities	0.3556	4.0164	−0.1564	3.9996
Communication	0.3149	4.6025	−0.2998	4.1980
Personal & Business Services	0.3746	6.1271	−0.3438	4.2622
Business Equipment	0.2820	7.7780	−0.6243	5.2615
Paper, Business Supplies, & Cont.	0.3527	5.0164	−0.4308	6.4425
Transportation	0.3922	5.7473	−0.5358	5.5043
Wholesale	0.2153	5.3179	−0.7175	6.6761
Retail	0.4704	5.5604	−0.5781	6.2862
Restaurants, Hotels, & Motels	0.4271	6.7383	−0.7662	5.9831
Banking, Insurance, RE, & Trading	0.5067	5.1682	−0.5156	5.0936
Everything Else	0.3370	5.5273	−0.2869	4.3823

Returns are continuously compounded in excess of the one-month Treasury bill yield (expressed in percent per month), and they are measured over the period from January 1960 to December 2005 ($T = 552$ months). Skewness is defined as the standardized third central moment of the data, and kurtosis is the standardized fourth central moment (Note: Kurtosis is not 'excess kurtosis' relative to the normal distribution.).

Table 6: Fitted Alphas from the Full Model: Industry Portfolios

Portfolio	α	StdErr(α)	t-statistic	Relative α
Food Products	0.1866	0.1090	1.7118	0.3513
Beer & Liquor	0.1079	0.1000	1.0796	0.1640
Tobacco Products	0.3134	0.1247	2.5142	0.3272
Recreation	0.1017	0.0953	1.0672	0.2830
Printing & Publishing	0.1048	0.1030	1.0182	0.2865
Consumer Goods	−0.0480	0.1022	−0.4695	−0.1169
Apparel	−0.0881	0.1102	−0.7994	−0.2026
Health Care, Med. Eq. & Pharm.	−0.0338	0.0888	−0.3802	−0.0650
Chemicals	−0.1218	0.0756	−1.6110	−0.4140
Textiles	−0.0480	0.1147	−0.4181	−0.1608
Construction & Materials	0.0180	0.0833	0.2162	0.0448
Steel Works Etc.	−0.0543	0.0816	−0.6655	−0.3512
Fabricated Products & Machinery	−0.0430	0.0913	−0.4710	−0.1137
Electrical Equipment	0.0930	0.1000	0.9299	0.2558
Automobiles & Trucks	0.0743	0.1294	0.5741	0.3406
Aircraft, Ships, & Railroad Eq.	0.0836	0.1232	0.6784	0.2013
Precious Metals & Mining	−0.1297	0.1394	−0.9305	−0.3131
Coal	−0.0100	0.1428	−0.0702	−0.0189
Petroleum & Natural Gas	0.1738	0.1029	1.6897	0.3156
Utilities	0.0203	0.0856	0.2366	0.0542
Communication	0.0964	0.1123	0.8583	0.2919
Personal & Business Services	−0.0576	0.1052	−0.5473	−0.1445
Business Equipment	−0.1750	0.1315	−1.3308	−0.5572
Paper, Business Supplies, & Cont.	−0.1322	0.0969	−1.3652	−0.3446
Transportation	−0.0169	0.0954	−0.1774	−0.0408
Wholesale	−0.1522	0.1061	−1.4344	−0.6257
Retail	0.1226	0.1106	1.1089	0.2471
Restaurants, Hotels, & Motels	0.0648	0.1218	0.5318	0.1398
Banking, Insurance, RE, & Trading	−0.0935	0.0977	−0.9570	−0.1757
Everything Else	−0.1163	0.1103	−1.0543	−0.3109

Alphas are computed using the standard two-pass approach applied to the conditional model with a second-order approximation to the risk premium function; i.e., using equations (24) and (25) in the text. They are expressed in percent per month. The data are from the period from January 1960 to December 2005 ($T = 552$ months). Standard errors for the alphas are constructed using equation (16) in the text. The t-statistic is the ratio of the estimate of alpha to its standard error. Relative alpha for a given portfolio is defined as the portfolio's alpha divided by the average excess return to the portfolio.

Table 7: Misspecification Alpha and Its Components: Industry Portfolios

Portfolio	α	α Components (a)	(b)	(c)	(d)	(e)	Rel. α
Food Products	0.142	0.062	−0.113	−0.564	0.775	−0.017	0.376
Beer & Liquor	0.303	0.231	−0.131	−0.683	0.942	−0.056	0.551
Tobacco Products	0.388	−0.040	−0.077	−0.246	0.674	0.076	0.601
Recreation	−0.023	−0.212	−0.162	−1.149	1.129	0.371	−0.091
Printing & Publishing	0.195	−0.058	−0.159	−0.806	0.991	0.227	0.748
Consumer Goods	0.294	0.058	−0.128	−0.553	0.838	0.080	0.641
Apparel	−0.204	0.029	−0.160	−1.089	0.924	0.093	−0.390
Health Care, Med. Eq. & Pharm.	0.276	0.119	−0.113	−0.582	0.876	−0.025	0.499
Chemicals	0.261	0.010	−0.135	−0.637	0.939	0.084	0.627
Textiles	0.004	−0.096	−0.136	−0.928	0.843	0.321	0.011
Construction & Materials	0.269	−0.121	−0.161	−0.761	1.075	0.237	0.700
Steel Works Etc.	0.390	−0.095	−0.163	−0.695	1.073	0.270	1.866
Fabricated Products & Machinery	0.372	−0.042	−0.148	−0.813	1.095	0.280	0.883
Electrical Equipment	0.241	−0.141	−0.126	−0.712	0.883	0.336	0.890
Automobiles & Trucks	0.013	0.012	−0.138	−0.899	0.839	0.199	0.089
Aircraft, Ships, & Railroad Eq.	0.265	−0.019	−0.153	−0.860	1.025	0.273	0.800
Precious Metals & Mining	0.352	−0.014	−0.094	−0.690	1.031	0.136	0.647
Coal	0.031	−0.133	−0.092	−0.587	0.830	0.013	0.058
Petroleum & Natural Gas	0.281	0.044	−0.088	−0.428	0.680	0.074	0.747
Utilities	0.107	0.046	−0.029	−0.372	0.407	0.055	0.303
Communication	0.134	−0.109	−0.066	−0.467	0.535	0.240	0.572
Personal & Business Services	0.142	−0.031	−0.138	−0.789	0.938	0.162	0.312
Business Equipment	0.463	0.007	−0.180	−0.860	1.188	0.308	0.947
Paper, Business Supplies, & Cont.	0.261	−0.014	−0.138	−0.651	0.994	0.069	0.505
Transportation	0.218	−0.165	−0.121	−0.686	0.848	0.342	0.505
Wholesale	0.132	−0.016	−0.122	−0.773	0.898	0.144	0.332
Retail	0.083	0.032	−0.160	−0.870	1.016	0.065	0.223
Restaurants, Hotels, & Motels	0.116	−0.026	−0.143	−1.093	1.012	0.365	0.290
Banking, Insurance, RE, & Trading	0.306	−0.125	−0.127	−0.691	0.990	0.258	0.489
Everything Else	0.295	0.034	−0.137	−0.618	0.910	0.105	0.601

This table shows the true (asymptotic) values of the portfolio alphas generated by the misspecification induced by the standard linear approximation. Overall alpha (and its components) are quoted in percent per month. They are constructed using equation (35) in the text, along with the parameter values in Table 8. The total alpha value is the sum of the different components:

$$
\begin{aligned}
(a) &= \boldsymbol{\lambda}_1' \left(\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} \right)^{-1} \Sigma_{\mathbf{F}_1\mathbf{R}^e} \\
(b) &= -\boldsymbol{\lambda}_1'\Sigma_{11}\Sigma_{12} \left(\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} \right)^{-1} \Sigma_{\mathbf{F}_2\mathbf{R}^e} \\
(c) &= -\boldsymbol{\lambda}_2'\Sigma_{22}\Sigma_{21} \left(\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} \right)^{-1} \Sigma_{\mathbf{F}_1\mathbf{R}^e} \\
(d) &= +\boldsymbol{\lambda}_2' \left(\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} \right)^{-1} \Sigma_{\mathbf{F}_2\mathbf{R}^e} \\
(e) &= -\widetilde{\boldsymbol{\lambda}}_1'\Sigma_{11}^{-1}\Sigma_{\mathbf{F}_1\mathbf{R}^e}
\end{aligned}
$$

Relative alpha is defined as true alpha divided by average excess returns net of the component of returns that the model cannot explain (α from Table 7).

Table 8: Simulation Results for an Economy Calibrated to Size and Book-to-Market Sorted Portfolios

Portfolio	Excess Returns Mean	Excess Returns StdDev	α Absolute Mean	α Absolute StdDev	α Relative Mean	α Relative StdDev
S1 BM1	0.1624	7.8157	−0.1234	0.1161	−0.5555	20.704
S1 BM2	0.3705	6.5957	−0.1083	0.1213	−0.4782	6.5230
S1 BM3	0.6411	5.6439	0.0042	0.1204	0.0078	0.2323
S1 BM4	0.8917	5.2601	0.1437	0.1230	0.1677	0.1508
S1 BM5	0.9796	5.5601	0.0830	0.1153	0.0872	0.1233
S2 BM1	0.1587	7.0123	−0.1448	0.1191	−1.9135	160.64
S2 BM2	0.5870	5.7002	0.0002	0.1154	−0.0062	0.5685
S2 BM3	0.6529	5.0306	0.0425	0.1155	0.0717	0.2043
S2 BM4	0.7778	4.8655	0.0834	0.1172	0.1128	0.1645
S2 BM5	1.0072	5.4646	0.0824	0.1187	0.0849	0.1249
S3 BM1	0.2176	6.4285	−0.0843	0.1129	−0.4324	8.5788
S3 BM2	0.5548	5.0855	0.0373	0.1106	0.0746	0.3614
S3 BM3	0.6580	4.6140	0.0481	0.1103	0.0775	0.1894
S3 BM4	0.7077	4.5142	0.0254	0.1100	0.0381	0.1695
S3 BM5	0.9112	5.1386	0.1236	0.1199	0.1396	0.1405
S4 BM1	0.2423	5.6727	−0.0244	0.1015	−0.1546	7.9475
S4 BM2	0.4905	4.8115	−0.0111	0.1009	−0.0255	0.5332
S4 BM3	0.4422	4.5680	−0.0884	0.1076	−0.2614	0.8109
S4 BM4	0.6562	4.4730	−0.0010	0.1060	−0.0018	0.1814
S4 BM5	0.7551	5.1123	0.0365	0.1176	0.0511	0.1709
S5 BM1	0.3593	4.5685	0.0828	0.0923	−0.0639	22.090
S5 BM2	0.4746	4.3112	0.0140	0.0921	0.0044	1.566
S5 BM3	0.5232	4.0905	0.0515	0.1007	0.1073	0.2311
S5 BM4	0.5206	4.0491	−0.0307	0.1073	−0.0638	0.2383
S5 BM5	0.6259	4.6156	−0.0658	0.1107	−0.1114	0.1964

This table reports average excess returns, average standard deviation of excess returns, and alphas – both absolute and relative – from 5,000 simulations of length $T = 550$ months of the model economy in Section 2 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 nonlinear model. OLS is used in fitting the second-pass cross-sectional regression. True alphas are zero. Alphas due to misspecification based on the true moments of the data are shown in the first column of Table 4.

Table 9: Simulation Results for an Economy Calibrated to Industry Portfolios

| | Excess Returns | | α Absolute | | Relative | |
Portfolio	Mean	StdDev	Mean	StdDev	Mean	StdDev
Food Products	0.3821	4.2795	0.0184	0.0681	0.0599	0.3779
Beer & Liquor	0.5556	5.1355	0.1451	0.0819	0.2759	0.1747
Tobacco Products	0.6430	6.7156	0.1345	0.0826	0.2183	0.1452
Recreation	0.2666	6.9414	−0.1092	0.0781	−0.7013	23.255
Printing & Publishing	0.2688	5.3725	−0.1048	0.0709	−0.0094	23.790
Consumer Goods	0.4624	4.6955	0.0617	0.0715	0.1521	0.2275
Apparel	0.5346	6.2831	0.0512	0.0788	0.1187	0.4179
Health Care, Med. Eq. & Pharm.	0.5588	4.8772	0.1151	0.0741	0.2258	0.1737
Chemicals	0.4227	5.0714	−0.0230	0.0692	−0.1139	3.6844
Textiles	0.3541	5.8041	0.0564	0.0735	0.2719	3.8590
Construction & Materials	0.3895	5.2269	−0.0445	0.0632	−0.8901	92.708
Steel Works Etc.	0.2123	6.4542	−0.1666	0.0817	−2.5347	109.24
Fabricated Products & Machinery	0.4254	5.6870	−0.0197	0.0670	−0.0831	1.9240
Electrical Equipment	0.2762	5.3817	−0.0838	0.0686	−1.0469	64.183
Automobiles & Trucks	0.1527	5.7488	−0.1140	0.0785	−0.3590	39.084
Aircraft, Ships, & Railroad Eq.	0.3394	6.2121	−0.0348	0.0771	−0.4155	21.982
Precious Metals & Mining	0.5496	6.8808	0.1245	0.0869	−0.2430	0.1942
Coal	0.5458	8.1832	0.0594	0.0802	−0.1160	0.1665
Petroleum & Natural Gas	0.3825	5.0049	−0.0053	0.0762	−0.0176	0.3541
Utilities	0.3590	3.9500	0.0977	0.0740	−0.2918	0.2493
Communication	0.2373	4.5027	−0.0551	0.0692	−0.6904	22.191
Personal & Business Services	0.4603	5.9094	−0.0171	0.0756	−0.0420	0.9017
Business Equipment	0.4944	7.5094	−0.0781	0.0766	−0.2418	2.7341
Paper, Business Supplies, & Cont.	0.5222	4.8017	0.0800	0.0672	0.1739	0.2309
Transportation	0.4381	5.5419	0.0117	0.0728	0.0569	2.9370
Wholesale	0.4007	5.0756	−0.0024	0.0650	−0.0052	1.1957
Retail	0.3817	5.2893	−0.0410	0.0700	−0.1054	1.7760
Restaurants, Hotels, & Motels	0.4065	6.3999	0.0650	0.0751	0.1671	5.6160
Banking, Insurance, RE, & Trading	0.6305	4.9333	0.1782	0.0684	0.3106	0.1760
Everything Else	0.4960	5.2994	−0.0001	0.0679	−0.0041	0.4065

This table reports average excess returns, average standard deviation of excess returns, and alphas – both absolute and relative – from 5,000 simulations of length $T = 550$ months of the model economy in Section 2 calibrated to industry 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 nonlinear model. OLS is used in fitting the second-pass cross-sectional regression. True alphas are zero. Alphas due to misspecification based on the true moments of the data are shown in the first column of Table 7.

Table 10: Rejection Rates for the χ^2-test for Nonlinear Risk Premium Terms in Simulations

Panel A: Testing without the Shanken Correction

	Nominal Size			
	0.100	0.050	0.025	0.010
Characteristics-Sorted Returns	0.980	0.968	0.953	0.927
Industry Returns	0.016	0.004	0.001	0.000

Panel B: Testing with the Shanken Correction

	Nominal Size			
	0.100	0.050	0.025	0.010
Characteristics-Sorted Returns	0.207	0.110	0.055	0.022
Industry Returns	0.000	0.000	0.000	0.000

The asymptotic distribution of the test for nonlinear risk premium terms in (19) and (20) is χ^2_6. The risk premiums are estimated using OLS in the second-pass cross-sectional regression. The critical values associated with the different nominal sizes are

Nominal Size	0.100	0.050	0.025	0.010
Critical Value	10.4664	12.5916	14.4494	16.8119

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 model economy in Section 2, calibrated to either characteristic-sorted returns or industry returns. Panel A examines the test statistic without a correction for stochastic betas (using $\widehat{\Sigma}_{\boldsymbol{\lambda}}$ of equation 17), while Panel B examines the test statistic that incorporates the Shanken correction ($\widetilde{\Sigma}_{\boldsymbol{\lambda}}$ of equation 18)

A   Parameters for Simulations

In this appendix, we report the parameter values used in computing estimates of the misspecification alphas and the simulations for the two different sets of test assets. The factor variances and covariances are the same for both sets of simulations, since the factors are the same.

There are two points worth noting about the calibrated risk premium parameters. First, there are differences – in both calibrations – between λ_1 and λ_1, the risk premium estimates on the first-order factors including and excluding the higher order terms. Also, the point estimates of the risk premiums on the omitted factors are large, both statistically (even using OLS standard errors) and relative to the risk premiums on the included factors. These facts are consistent with a large value for alpha due to model misspecification.

Table A-1: Calibrated Values of Model Parameters: size and B/M Sorted Returns

Panel A: Factor Covariance Matrices

$$\Sigma_{11} = \begin{bmatrix} 1.000 & & & & & & \\ 0.000 & 1.000 & & & & & \\ 0.000 & 0.000 & 1.000 & & & & \\ 0.006 & -0.044 & 0.121 & 0.998 & & & \\ -0.236 & 0.045 & -0.230 & 0.100 & 1.088 & & \\ 0.045 & -0.291 & -0.016 & -0.025 & -0.035 & 1.266 & \\ -0.230 & -0.017 & -0.095 & 0.027 & 0.273 & 0.039 & 0.862 \end{bmatrix}$$

$$\Sigma_{22} = \begin{bmatrix} 1.000 & & & \\ 0.263 & 1.055 & & \\ 0.092 & 0.178 & 1.733 & \\ 0.062 & 0.162 & -0.193 & 0.823 \end{bmatrix}$$

$$\Sigma_{12} = \begin{bmatrix} 0.069 & 0.052 & -0.022 & 0.073 \\ -0.019 & -0.022 & 0.173 & 0.025 \\ 0.007 & 0.071 & 0.025 & -0.033 \\ -0.002 & -0.080 & -0.077 & 0.164 \\ -0.080 & -0.260 & 0.055 & -0.176 \\ -0.077 & 0.058 & -0.543 & -0.050 \\ 0.164 & -0.184 & -0.048 & -0.159 \end{bmatrix}$$

Panel B: Risk Premium Estimates from size and B/M-Sorted Returns

$$\lambda_1' = \begin{bmatrix} -0.524 & -2.401 & 0.163 & 0.644 & 0.229 & -0.581 & 0.344 \\ (0.125) & (0.169) & (0.053) & (0.176) & (0.091) & (0.099) & (0.110) \end{bmatrix}'$$

$$\lambda_2' = \begin{bmatrix} -0.931 & -0.347 & -0.361 & 0.267 \\ (0.134) & (0.152) & (0.237) & (0.143) \end{bmatrix}'$$

$$\widetilde{\lambda}_1' = \begin{bmatrix} -0.053 & -2.038 & 0.099 & 0.987 & 0.217 & -0.316 & 0.073 \\ (0.071) & (0.103) & (0.047) & (0.115) & (0.086) & (0.081) & (0.082) \end{bmatrix}'$$

Panel C: Risk Premium Estimates from Industry Returns

$$\lambda_1' = \begin{bmatrix} -0.240 & -0.343 & 0.084 & -0.024 & 0.183 & -0.349 & -0.093 \\ (0.047) & (0.050) & (0.042) & (0.049) & (0.055) & (0.061) & (0.045) \end{bmatrix}'$$

$$\lambda_2' = \begin{bmatrix} -0.134 & -0.257 & 0.972 & 0.070 \\ (0.049) & (0.053) & (0.065) & (0.044) \end{bmatrix}'$$

$$\widetilde{\lambda}_1' = \begin{bmatrix} -0.311 & -0.310 & 0.043 & -0.128 & -0.174 & -0.656 & -0.159 \\ (0.045) & (0.049) & (0.041) & (0.047) & (0.052) & (0.053) & (0.043) \end{bmatrix}'$$

Σ_{11}, Σ_{22}, and Σ_{12} are the point estimates of the covariance matrix of the first-order factors, the second-order factors, and their covariance, respectively. λ_1 and λ_2 are risk premium estimates (standard errors in parentheses) from the complete second-order estimation. $\widetilde{\lambda}_1$ is the estimate of the first-order risk premiums (standard errors in parentheses) from the model that omits higher-order terms.

References

[1] 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.

[2] Bekaert, G., and J. Lui, 2004, "Conditioning Information and Variance Bounds on Asset Pricing Kernels," *Review of Financial Studies 17*, pages 339-378.

[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 D. A. Chapman, 2005, "Comparing Multifactor Models of the Term Structure," Manuscript, Duke University and Boston College.

[6] 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.

[7] 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.

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

[9] 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.

[10] Chan, K. C., G. A. Karolyi, F. A. Longstaff, and A. B. Sanders, 1992, "An Empirical Comparison of Alternative Models of the Short-Term Interest Rate," *Journal of Finance, 47*, pages 1209-1227.

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

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

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

[14] Dai, Q., and K. J. Singleton, 2000, "Specification Analysis of Affine Term Structure Models," *Journal of Finance, 55*, pages 1943-1978.

[15] Dai, Q., K. J. Singleton, and W. Yang, 2005, "Regime Shifts in a Dynamic Term Structure Model of U.S. Treasury Bond Yields," Manuscript, Stanford University.

[16] Daniel, K., and S. Titman, 2005, "Testing Factor-Model Explanations of Market Anomalies," Manuscript, Northwestern University and University of Texas at Austin.

[17] Dittmar, R. F., 2002, "Nonlinear Pricing Kernels, Kurtosis Preference, and Evidence from the Cross-Section of Equity Returns," *Journal of Finance, 57*, pages 369-403.

[18] Duffee, G. R., 2002, "Term Premia and Interest Rate Forecasts in Affine Models," *Journal of Finance, 57*, pages 405-443.

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

[20] 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.

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

[22] 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.

[23] Ferson, W. E., and A. F. Siegel, 2001, "The efficient use of Conditioning Information in Portfolios," *Journal of Finance 56*, pages 967-982.

[24] Gallant, A. R., L. P. Hansen, and G. Tauchen, 1990, "Using Conditional Moments of Asset Payoffs to Infer the Volatility of Intertemporal Marginal Rates of Substitution," *Journal of Econometrics, 45*, pages 141-179.

[25] 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.

[26] 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.

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

[28] Hansen, L. P., 1982, "Large Sample Properties of General Method of Moment Estimators," *Econometrica, 50,* pages 1029-1054.

[29] 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.

[30] 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.

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

[32] 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.

[33] 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.

[34] Lochstoer, L., 2006, "Expected Returns and the Business Cycle: Heterogeneous Agents and Heterogeneous Goods," Manuscript, London Business School.

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

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

[37] Phalippou, L., 2005, "Institutional Ownership and the Value Premium," Manuscript, University of Amsterdam.

[38] Santos, T., and P. Veronesi, 2005, "Labor Income and Predictable Stock Returns," forthcoming in *Review of Financial Studies*.

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

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