Robust Inference in Linear Asset Pricing Models

NIKOLAY GOSPODINOV, RAYMOND KAN, and CESARE ROBOTTI*

First draft: August 2011
This draft: March 2013

JEL # G12, C13, C32

*Gospodinov is from Concordia University and CIREQ. Kan is from the University of Toronto. Robotti is from the Federal Reserve Bank of Atlanta and EDHEC Risk Institute. We thank Chu Zhang, Guofu Zhou, seminar participants at the Federal Reserve Bank of Atlanta, Imperial College London, Singapore Management University, and University of Georgia, and participants at the 3rd Annual CIRPÉE Applied Financial Time Series Workshop at HEC (Montreal) and Mathematical Finance Days 2012. Gospodinov gratefully acknowledges financial support from Fonds de recherche sur la société et la culture (FQRSC), Institut de Finance Mathématique de Montréal (IFM2), and the Social Sciences and Humanities Research Council of Canada. Kan gratefully acknowledges financial support from the Social Sciences and Humanities Research Council of Canada and the National Bank Financial of Canada. The views expressed here are the authors' and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System. Corresponding author: Cesare Robotti, Research Department, Federal Reserve Bank of Atlanta, 1000 Peachtree Street NE, Atlanta, Georgia, 30309-4470; E-mail: cesare.robotti@atl.frb.org.
Robust Inference in Linear Asset Pricing Models

Abstract

Many asset pricing models include risk factors that are only weakly correlated with the asset returns. We show that in the presence of a factor that is independent of the returns ("useless factor"), the standard inference procedures for evaluating its pricing ability could be highly misleading in misspecified models. Our proposed model selection procedure, which is robust to useless factors and potential model misspecification, restores the standard inference and proves to be effective in eliminating factors that do not improve the model’s pricing ability. The practical relevance of our analysis is illustrated using simulations and an empirical application.
Misspecification is an inherent feature of many asset pricing models and reliable statistical inference crucially depends on its robustness to potential model misspecification. Kan and Robotti (2008, 2009) and Gospodinov, Kan, and Robotti (2013) show that by ignoring model misspecification, one can mistakenly conclude that a risk factor is priced when, in fact, it does not contribute to the pricing ability of the model. The problem can be particularly serious when the pricing model includes factors that are weakly correlated with the returns on the test assets, such as macroeconomic factors. While the three papers mentioned above provide a general statistical framework for inference, evaluation and comparison of potentially misspecified asset pricing models, the misspecification-robust inference in these papers is developed under the assumption that the covariance matrix of asset returns and risk factors is of full column rank.

In the extreme case of model misspecification with one or more “useless” factors (i.e., factors that are independent of the asset returns), the identification condition fails (i.e., the covariance matrix of asset returns and risk factors is of reduced rank) and the validity of the statistical inference is compromised. The impact of the violation of this identification condition on the asymptotic properties of parameter hypothesis and specification tests in models estimated via two-pass cross-sectional regressions and generalized method of moments (GMM) was first studied by Kan and Zhang (1999a, 1999b). Burnside (2010, 2011) discusses analogous identification failures for alternative normalizations of the stochastic discount factor (SDF). Kleibergen (2009, 2010) and Khalaf and Schaller (2011) propose test procedures that exhibit robustness to the degree of correlation between returns and factors in a two-pass cross-sectional regression framework.

In this paper, we focus on linear SDFs mainly because the useless factor problem is well-defined for this class of models. In addition, we choose to present our results for the distance metric introduced by Hansen and Jagannathan (HJ, 1997). This measure has gained tremendous popularity in the empirical asset pricing literature and has been used both as a model diagnostic and as a tool for model selection by many researchers.1 In particular, we investigate whether the misspecification-robust standard errors proposed by Hall and Inoue (2003) and Kan and Robotti (2008, 2009) can guard the standard inference against the presence of useless factors. The main contributions of our analysis can be summarized as follows. First, we demonstrate that the misspecification-robust

---

1While we study explicitly only the GMM estimator based on the HJ-distance, our results continue to hold for the class of optimal GMM estimators. Some simulation results for the optimal GMM case are provided in an online appendix available on the authors’ websites.
Wald test for the useless factor is asymptotically distributed as a chi-squared random variable with one degree of freedom. This stands in sharp contrast with the Wald test constructed under the assumption of correct specification which is shown to be asymptotically chi-squared distributed with degrees of freedom given by the difference between the number of assets and the number of (useful and useless) factors included in the model. An immediate implication of the latter result is that using standard inference will result in substantial over-rejection of the null hypothesis that the risk premium on the useless factor is equal to zero.\(^2\)

Second, we show that the estimator of the coefficient associated with the useless factor diverges with the sample size while the remaining parameters are not consistently estimable. The limiting distributions of the \(t\)-statistics corresponding to the useful factors are found to be non-standard and less dispersed when a useless factor is present. Regardless of whether the model is correctly specified or misspecified, the misspecification-robust standard errors ensure asymptotically valid inference and allow us to identify factors that do not contribute to the pricing of the test assets (i.e., useless factors and factors that do not reduce the HJ-distance).

Third, we show that the specification test based on the HJ-distance is inconsistent in the presence of a useless factor. To restore the standard inference for the \(t\)-tests on the parameters associated with the useful factors and for the test of correct model specification, we propose a simple sequential procedure which allows us to eliminate the useless factors from the model. Monte Carlo simulation results suggest that our sequential model selection procedure is effective in retaining useful factors in the model and eliminating factors that are either useless or do not reduce the HJ-distance. As a result, our proposed method can guard against both model misspecification and the presence of useless factors in the analysis.

Empirically, our interest is in robust estimation of several prominent asset pricing models with macroeconomic and financial factors using the HJ-distance measure. In addition to the basic CAPM and consumption CAPM (CCAPM), the theory-based models considered in our main empirical analysis are the CAPM with labor income of Jagannathan and Wang (1996), the CCAPM conditioned on the consumption-wealth ratio of Lettau and Ludvigson (2001), the durable consumption model (D-CCAPM) of Yogo (2006), and the five-factor implementation of the intertemporal CAPM

\(^2\)Our use of the term “over-rejection” is somewhat non-standard since the true risk premium on a useless factor is not identifiable. Nevertheless, since a useless factor does not improve the pricing performance of the model, testing the null of a zero risk premium is of most practical importance.
used by Petkova (2006). We also study the well-known “three-factor model” of Fama and French (1993). Although this model was primarily motivated by empirical observation, its size and book-to-market factors are sometimes viewed as proxies for more fundamental economic factors.

Our main empirical analysis uses the monthly returns on the 25 size and book-to-market portfolios of Fama and French (1993) and the one-month T-bill from February 1959 until July 2007. The HJ-distance test rejects the hypothesis of a perfect fit for all models except for the ICAPM. In addition, the test for reduced rank indicates that only the CAPM and the three-factor model of Fama and French (1993), two models with traded factors only, are properly identified. This clearly points to the need for statistical methods that are robust to model misspecification and weak identification. We show empirically that when misspecification-robust standard errors are employed, only the market, book-to-market, term premium, and nondurable consumption growth factors survive our sequential procedure at the 5% significance level.

It is important to stress that the useless factor problem is not an isolated problem limited to the data and asset pricing models considered in our main empirical analysis. We show that qualitatively similar pricing conclusions can be reached using different data frequencies and SDF specifications. Overall, our results suggest that the statistical evidence on the pricing ability of many macroeconomic and financial factors is weak and their usefulness in explaining the cross-section of asset returns should be interpreted with caution.

The rest of the paper is organized as follows. Section 1 reviews some of the main results for asymptotically valid inference under potential model misspecification. In Section 2, we introduce a useless factor in the analysis and present the limiting distributions of the parameters of interest and their $t$-statistics under both correct model specification and model misspecification. We also discuss some practical implications of our theoretical analysis and suggest an easy-to-implement model selection procedure. Section 3 reports results from a Monte Carlo simulation experiment. In Section 4, we conduct an empirical investigation of some popular asset pricing models with traded and non-traded factors. Section 5 concludes.
1. Asymptotic Inference with Useful Factors

This section introduces the notation and reviews some main results that will be used in the subsequent analysis. Let

\[ y_t(\gamma_1) = \tilde{f}_t\gamma_1 \]  

(1)

be a candidate linear SDF, where \( \tilde{f}_t = [1, f_t'] \) is a \( K \)-vector with \( f_t \) being a \((K - 1)\)-vector of risk factors, and \( \gamma_1 \) is a \( K \)-vector of SDF parameters with generic element \( \gamma_{1i} \) for \( i = 1, \ldots, K \). Also, let \( x_t \) be the random payoffs of \( N \) assets at time \( t \) and \( q \neq 0_N \) be a vector of their original costs.\(^3\)

We assume throughout that the second moment matrix of \( x_t \), \( U = E[x_t x_t'] \), is nonsingular so that none of the test assets is redundant. In addition, we assume \( B = E[x_t \tilde{f}_t] \) is of full column rank but this assumption will be relaxed later on when we introduce a useless factor into the model. The specification in (1) is general enough to allow \( \tilde{f}_t \) to include cross-product factor terms (using lagged state variables as scaling factors).

Define the model pricing errors as

\[ e(\gamma_1) = E[x_t \tilde{f}_t\gamma_1 - q] = B\gamma_1 - q. \]  

(2)

If there exists no value of \( \gamma_1 \) for which \( e(\gamma_1) = 0_N \), the model is misspecified. This corresponds to the case when \( q \) is not in the span of the column space of \( B \). The pseudo-true parameter vector \( \gamma_1^* \) is defined as the solution to the quadratic minimization problem

\[ \gamma_1^* = \arg \min_{\gamma_1 \in \Gamma_1} e(\gamma_1)' W e(\gamma_1) \]  

(3)

for some symmetric and positive-definite weighting matrix \( W \), where \( \Gamma_1 \) denotes the parameter space.

The HJ-distance is obtained when \( W = U^{-1} \) and is given by

\[ \delta = \sqrt{e(\gamma_1^*)' U^{-1} e(\gamma_1^*)}. \]  

(4)

Given the computational simplicity and the nice economic and maximum pricing error interpretation of the HJ-distance, this measure of model misspecification is widely used in applied work for

\(^3\)When \( q = 0_N \), i.e., when only excess returns are used, the mean of the SDF cannot be identified and researchers have to choose some normalization of the SDF (see, for example, Kan and Robotti, 2008, and Burnside, 2010). Theoretical and simulation results for the analysis with excess returns are provided in an online appendix available on the authors' websites.
estimation and evaluation of asset pricing models. For this reason, we consider explicitly only the case of the HJ-distance although results for the optimal GMM estimator are also available from the authors upon request.

The estimator \( \hat{\gamma}_1 \) of \( \gamma^*_1 \) is obtained by minimizing the sample analog of (3):

\[
\hat{\gamma}_1 = \arg \min_{\gamma_1 \in \Gamma_1} \hat{e}(\gamma_1)'\hat{U}^{-1}\hat{e}(\gamma_1),
\]

where \( \hat{U} = \frac{1}{T} \sum_{t=1}^{T} x_t x'_t \), \( \hat{e}(\gamma_1) = \hat{B}\gamma_1 - q \) and

\[
\hat{B} = \frac{1}{T} \sum_{t=1}^{T} x_t \tilde{f}_t.
\]

Then, the solution to the above minimization problem is given by

\[
\hat{\gamma}_1 = (\hat{B}'\hat{U}^{-1}\hat{B})^{-1}\hat{B}'\hat{U}^{-1}q.
\]

Let \( e_t(\gamma^*_1) = x_t \tilde{f}_t \gamma^*_1 - q \) and \( S = E[e_t(\gamma^*_1)e_t(\gamma^*_1)'] \). For ease of exposition, we assume that \( e_t(\gamma^*_1) - e(\gamma^*_1) \) forms a stationary and ergodic martingale difference sequence. Under suitable regularity conditions, Kan and Robotti (2009) show that

\[
\sqrt{T}(\hat{\gamma}_1 - \gamma^*_1) \overset{d}{\to} N(0, \Sigma_{\hat{\gamma}_1}),
\]

where \( \Sigma_{\hat{\gamma}_1} = E[h_t h'_t] \),

\[
h_t = (B'U^{-1}B)^{-1}B'U^{-1}e_t(\gamma^*_1) + (B'U^{-1}B)^{-1}(\tilde{f}_t - B'U^{-1}x_t)u_t
\]

and

\[
u_t = e(\gamma^*_1)'U^{-1}x_t.
\]

Note that if the model is correctly specified (i.e., \( u_t = 0 \)), the expression for \( h_t \) specializes to

\[
h_t^0 = (B'U^{-1}B)^{-1}B'U^{-1}e_t(\gamma^*_1)
\]

and the asymptotic covariance matrix of \( \sqrt{T}(\hat{\gamma}_1 - \gamma^*_1) \) is simplified to

\[
\Sigma_{\hat{\gamma}_1}^0 = E[h_t^0 h_t'^0] = (B'U^{-1}B)^{-1}B'U^{-1}SU^{-1}B(B'U^{-1}B)^{-1}.
\]

Suppose now that the interest lies in testing hypotheses on the individual parameters of the form \( H_0 : \gamma_{1i} = \gamma^*_{1i} \) (for \( i = 1, \ldots, K \)) and define a selector vector \( \iota_i \) with one for its \( i \)-th element and zero
otherwise (the length of $i$ is implied by the matrix that it is multiplied to). Then, the $t$-statistic for $\tilde{\gamma}_{1i}$ with standard error computed under potential model misspecification is asymptotically distributed as

$$t_m(\tilde{\gamma}_{1i}) = \frac{\tilde{\gamma}_{1i} - \gamma^*_1}{\sqrt{i'\hat{\Sigma}_{1i}i}} \xrightarrow{d} N(0,1),$$

where $\hat{\Sigma}_{1i}$ is a consistent estimator of $\Sigma_{1i}$. Note that this result is valid irrespective of whether the model is misspecified or correctly specified.

In applied work, it is a common practice to test parameter restrictions using $t$-tests based on standard errors computed under the assumption of correct model specification. For this reason, it is instructive to consider the large sample behavior of the $t$-test

$$t_c(\tilde{\gamma}_{1i}) = \frac{\tilde{\gamma}_{1i} - \gamma^*_1}{\sqrt{i'\hat{\Sigma}_{1i}i}},$$

where $\hat{\Sigma}_{1i}^0$ is a consistent estimator of $\Sigma_{1i}^0$. If the model is indeed correctly specified, the $t$-test $t_c(\tilde{\gamma}_{1i})$ is asymptotically distributed as a standard normal random variable

$$t_c(\tilde{\gamma}_{1i}) \xrightarrow{d} N(0,1).$$

However, using the result in (8)–(9), we have that under misspecified models

$$t_c(\tilde{\gamma}_{1i}) \xrightarrow{d} N\left(0, \frac{i'\hat{\Sigma}_{1i}i}{i'\hat{\Sigma}_{1i}^0i}\right).$$

Furthermore, under the assumption that $x_t$ and $f_t$ are multivariate elliptically distributed, it can be shown (Kan and Robotti, 2009) that $(i'\hat{\Sigma}_{1i}i)/(i'\hat{\Sigma}_{1i}^0i) > 1$, which implies that standard inference based on critical values from the $N(0,1)$ distribution would tend to over-reject the null hypothesis.

We conclude this section with several observations that emerge from a closer inspection of the function $h_t$ in (9) which is used for computing the covariance matrix $\Sigma_{1i}$ under misspecification. It proves useful to rewrite $h_t$ as

$$h_t = h_t^0 + (B'U^{-1}B)^{-1}(\tilde{f}_t - B'U^{-1}x_t)u_t.$$

The adjustment term $(B'U^{-1}B)^{-1}(\tilde{f}_t - B'U^{-1}x_t)u_t$ contains three components: (i) a misspecification component $u_t$, (ii) a spanning component $\tilde{f}_t - B'U^{-1}x_t$ that measures the degree to which the factors are mimicked by the returns on the test assets, and (iii) a component $(B'U^{-1}B)^{-1}$ that
measures the usefulness of factors. The adjustment term is zero if the model is correctly specified \((u_t = 0)\) and/or the factors are fully mimicked by the returns \((\tilde{f}_t = B'U^{-1}x_t)\). If the factors are nearly uncorrelated with the returns (i.e., \(B\) is close to zero), the component \((B'U^{-1}B)^{-1}\) can be very large and the adjustment term tends to dominate the behavior of \(h_t\).

2. Asymptotic Inference in the Presence of a Useless Factor

As argued in the introduction, many popular asset pricing models include macroeconomic risk factors that often have very low correlations with the returns on the test assets. For this reason, we now consider a candidate SDF which is given by

\[
y_t = \tilde{f}_t\gamma_1 + g_t\gamma_2,
\]

where \(g_t\) is assumed to be a useless factor such that it is independent of \(x_t\) and \(f_t\) for all time periods. For ease of exposition, we assume that \(E[g_t] = 0\) and \(\text{Var}[g_t] = 1\).\(^4\) Note that the independence between \(g_t\) and \(x_t\) implies

\[
d = E[x_tg_t] = 0_N
\]

and

\[
E[x_t'x_tg_t^2] = E[E[x_t'x_t|g_t]g_t^2] = UE[g_t^2] = U.
\]

Now let \(D = [B, d], \gamma = [\gamma_1', \gamma_2]', e(\gamma) = D\gamma - q, \hat{d} = \frac{1}{T} \sum_{t=1}^{T} x_tg_t\), and \(\hat{D} = [\hat{B}, \hat{d}]\). Note that since \(d = 0_N\), the vector of pricing errors

\[
e(\gamma) = B\gamma_1 + d\gamma_2 - q = B\gamma_1 - q
\]

is independent of the choice of \(\gamma_2\). For the pseudo-true values of the SDF parameters, we can set \(\gamma_1^*\) as in (3) but the parameters associated with the useless factor \((\gamma_2^*)\) cannot be identified. In the following, we set \(\gamma_2^* = 0\), which is a natural choice because in Theorem 1 we will show that \(\hat{\gamma}_2\) is symmetrically distributed around zero.

\(^4\)The independence of the useless factor from the test asset returns and the other factors is a sufficient condition for our results to go through. The assumption of zero mean for the useless factor does not affect our asymptotic results on statistical inference for the slope parameters of the linear SDF. It does, however, affect the limiting distribution of the estimated SDF’s intercept and the statistical inference on it. The limiting results derived under a generic mean and variance of the useless factor are available from the authors upon request.
While the pseudo-true values of $\gamma^*_2$ are not identified, the sample estimates of the SDF parameters are always identified and they are given by

$$\hat{\gamma} = (\hat{D}'\hat{U}^{-1}\hat{D})^{-1}\hat{D}'\hat{U}^{-1}q.$$  \hfill(22)

Note that this is equivalent to running an ordinary least squares (OLS) regression of $\hat{D}^{-1/2}q$ on $\hat{D}^{-1/2}\hat{B}$ and $\hat{D}^{-1/2}\hat{d}$. In order to obtain $\hat{\gamma}_2$, we can project $\hat{D}^{-1/2}q$ and $\hat{D}^{-1/2}\hat{d}$ on $\hat{D}^{-1/2}\hat{B}$, and then regress the residuals from the first projection onto the residuals from the second projection. It follows that

$$\hat{\gamma}_2 = \frac{\hat{d}'\hat{D}^{-1/2}(I_N - \hat{D}^{-1/2}\hat{B}(\hat{B}'\hat{D}^{-1/2}\hat{B})^{-1}\hat{B}'\hat{D}^{-1/2}q)}{\hat{d}'\hat{D}^{-1/2}(I_N - \hat{D}^{-1/2}\hat{B}(\hat{B}'\hat{D}^{-1/2}\hat{B})^{-1}\hat{B}'\hat{D}^{-1/2}d).}$$  \hfill(23)

Similarly, the parameter vector $\hat{\gamma}_1$ is obtained by projecting $\hat{D}^{-1/2}q$ and $\hat{D}^{-1/2}\hat{B}$ on $\hat{D}^{-1/2}\hat{d}$ and then regressing the residuals from the first projection onto the residuals from the second projection, which yields

$$\hat{\gamma}_1 = (\hat{B}'\hat{D}^{-1/2}(I_N - \hat{D}^{-1/2}\hat{d}(\hat{D}'\hat{D}^{-1/2}\hat{d})^{-1}\hat{D}'\hat{D}^{-1/2}q)^{-1}$$

$$\times \hat{B}'\hat{D}^{-1/2}(I_N - \hat{D}^{-1/2}\hat{d}(\hat{D}'\hat{D}^{-1/2}\hat{d})^{-1}\hat{D}'\hat{D}^{-1/2}q).$$  \hfill(24)

Our first results are concerned with the limiting behavior of $\hat{\gamma}_1$ and $\hat{\gamma}_2$ under correctly specified and misspecified models. We adopt the following notation. Let $\hat{B} = U^{-1/2}B$, $\hat{q} = U^{-1/2}q$, and $P$ be an $N \times (N - K)$ orthonormal matrix whose columns are orthogonal to $\hat{B}$ so that $PP' = I_N - \hat{B}(\hat{B}'\hat{B})^{-1}\hat{B}'$. Also, let $z \sim N(0_N, I_N)$ and $y \sim N(0_N, U^{-1/2}SU^{-1/2})$, and they are independent of each other. Finally, we define $w = P'z \sim N(0_{N-K}, I_{N-K})$, $s = (\hat{q}'Pw)/(\hat{q}'PP'\hat{q})^{1/2} \sim N(0, 1)$, $u = P'y \sim N(0_{N-K}, V_u)$ with $V_u = P'U^{-1/2}SU^{-1/2}P$, and $r = (\hat{B}'\hat{B})^{-1/2}\hat{B}'y \sim N(0_K, V_r)$ with $V_r = (\hat{B}'\hat{B})^{-1/2}\hat{B}'U^{-1/2}SU^{-1/2}\hat{B}(\hat{B}'\hat{B})^{-1/2}$.  

**Theorem 1.** Assume that $N > K + 1$, $[x_t, k_t, g_t]'$ are jointly stationary and ergodic processes with finite fourth moments, and $e_t(\gamma^*_2) - e(\gamma^*_2)$ forms a martingale difference sequence.

(a) If $\delta = 0$, i.e., the model is correctly specified, we have

$$\sqrt{T}(\hat{\gamma}_1 - \gamma^*_1) \xrightarrow{d} (\hat{B}'\hat{B})^{-1/2} \left[ r - \frac{w'u}{w'w} (\hat{B}'\hat{B})^{-1/2} \hat{B}'z \right],$$  \hfill(25)

and

$\hat{\gamma}_2 \xrightarrow{d} \frac{w'u}{w'w}$.  \hfill(26)
(b) If $\delta > 0$, i.e., the model is misspecified, we have

$$
\hat{\gamma}_1 - \gamma^*_1 \xrightarrow{d} - \frac{\delta s}{w'w}(\hat{B}'\hat{B})^{-1}\hat{B}'z,
$$

(27)

and

$$
\frac{1}{\sqrt{T}}\hat{\gamma}_2 \xrightarrow{d} \frac{\delta s}{w'w}.
$$

(28)

Proof. See the Appendix.

The results in Theorem 1 can be summarized as follows. First, for correctly specified models, Theorem 1 shows that $\hat{\gamma}_2$ converges to a bounded random variable rather than zero.\footnote{The limiting random variable has mean zero and variance $\text{tr}(V_u)/[(N - K)(N - K - 2)]$, where $\text{tr}(\cdot)$ is the trace operator.} While the parameter estimates for the useful factors are consistently estimable, they are asymptotically non-normally distributed. Second, the presence of a useless factor further exacerbates the inference problems when the model is misspecified. In this case, the estimator $\hat{\gamma}_1$ is inconsistent while the estimator $\hat{\gamma}_2$ diverges at rate $T^{1/2}$ which is in agreement with the results in Kan and Zhang (1999b) and Kleibergen (2009).

Despite the highly non-standard limits of the SDF parameter estimates, it is possible that their $t$-statistics are well behaved. To investigate this, we define two types of $t$-statistics: (i) $t_c(\hat{\gamma}_{1i})$, for $i = 1, \ldots, K$, and $t_c(\hat{\gamma}_2)$ that use standard errors obtained under the assumption that the model is correctly specified, and (ii) $t_m(\hat{\gamma}_{1i})$, for $i = 1, \ldots, K$, and $t_m(\hat{\gamma}_2)$ that use standard errors under potentially misspecified models. The two types of $t$-statistics are based on the estimated covariance matrices $\hat{\Sigma}_0 = \frac{1}{T} \sum_{t=1}^{T} \hat{h}_t^0\hat{h}_t^0$ and $\hat{\Sigma}_\delta = \frac{1}{T} \sum_{t=1}^{T} \hat{h}_t\hat{h}_t'$, where

$$
\hat{h}_t^0 = (\hat{D}'\hat{U}^{-1}\hat{D})^{-1}\hat{D}'\hat{U}^{-1}\hat{e}_t,
$$

(29)

$$
\hat{h}_t = \hat{h}_t^0 + (\hat{D}'\hat{U}^{-1}\hat{D})^{-1}(\hat{f}_t', g_t') - \hat{D}'\hat{U}^{-1}x_t)e'\hat{U}^{-1}x_t,
$$

(30)

$\hat{e}_t = x_t(\hat{f}_t\hat{\gamma}_1 + g_t\hat{\gamma}_2) - q$ and $\hat{e} = \frac{1}{T} \sum_{t=1}^{T} \hat{e}_t$. We explicitly consider the behavior of $t_c(\hat{\gamma}_{1i})$ and $t_c(\hat{\gamma}_2)$ because it is a common practice for researchers to assume correct specification when computing the $t$-statistics.

In particular, the $t$-statistics of $H_0 : \gamma_{1i} = \gamma^*_{1i}$ and $H_0 : \gamma_2 = 0$ under the assumption of a correctly specified model have the form

$$
t_c(\hat{\gamma}_{1i}) = \frac{\sqrt{T}(\hat{\gamma}_{1i} - \gamma^*_{1i})}{\sqrt{t_i^0\hat{\Sigma}_0 t_i}}
$$

(31)
and
\[ t_c(\hat{\gamma}_2) = \frac{\sqrt{T}\hat{\gamma}_2}{\sqrt{\hat{\gamma}_2'\hat{\Sigma}_0}\hat{\gamma}_2} \] (32)

Similarly, the \( t \)-statistics of \( H_0 : \gamma_{1i} = \gamma_{1i}^* \) and \( H_0 : \gamma_2 = 0 \) under a potentially misspecified model are given by
\[ t_m(\hat{\gamma}_{1i}) = \frac{\sqrt{T}(\hat{\gamma}_{1i} - \gamma_{1i}^*)}{\sqrt{\hat{\gamma}_{1i}'\hat{\Sigma}_{1i}\hat{\gamma}_{1i}}} \] (33)

and
\[ t_m(\hat{\gamma}_2) = \frac{\sqrt{T}\hat{\gamma}_2}{\sqrt{\hat{\gamma}_2'\hat{\Sigma}_{1i}\hat{\gamma}_2}} \] (34)

The results presented below are driven, to a large extent, by the limiting behavior of the matrix \( \hat{S} = \frac{1}{T}\sum_{t=1}^{T} \hat{e}_t\hat{e}_t' \). In the presence of a useless factor, the results in Theorem 1 imply that for misspecified models
\[ \hat{e}_t = (T^{-\frac{1}{2}}\hat{\gamma}_2)(T^{\frac{1}{2}}x_t g_t) + O_p(1) \] (35)

and
\[ \frac{\hat{S}}{T} = (T^{-\frac{1}{2}}\hat{\gamma}_2)^2U + o_p(1), \] (36)

so \( \hat{S} \) diverges at rate \( T \). In contrast, for correctly specified models, we have
\[ \hat{S} = S + \hat{\gamma}_2^2U + o_p(1) \] (37)

so that \( \hat{S} \) converges to a random matrix.

In addition to the random variables and matrices defined before Theorem 1, we introduce the following notation. Let \( \tilde{u} \sim N(0,1), \tilde{r}_i \sim N(0,1), \tilde{z}_i \sim N(0,1), v \sim \chi^2_{N-K-1} \), and they are independent of each other and \( w \). Theorem 2 and Corollary 1 below provide the limiting distributions of the \( t \)-statistics under correctly specified and misspecified models.

**Theorem 2.**

(a) Suppose that the assumptions in Theorem 1 hold. In addition, assume that \( E[\epsilon_t\epsilon_t'|\tilde{f}_t] = \Sigma \) (conditional homoskedasticity), where \( \epsilon_t = x_t - B(E[\tilde{f}_t\tilde{f}_t'^{-1}]\tilde{f}_t) \). If \( \delta = 0 \), i.e., the model is

\[ \text{The limiting distribution of } t_c(\hat{\gamma}_2) \text{ does not depend on the conditional homoskedasticity assumption. The expressions for the limiting distributions of the other } t \text{-statistics under conditional heteroskedasticity are more involved, and the results are available upon request.} \]
correctly specified, we have

\[
\begin{align*}
t_c(\hat{\gamma}_{1i}) & \xrightarrow{d} \frac{\tilde{u}\tilde{z}_i + \sqrt{\lambda_i} \sqrt{w'w} \tilde{r}_i}{\left[\lambda_i w'w + \tilde{z}_i^2 + \tilde{u}^2 \left(1 + \frac{\tilde{z}_i^2}{w'w}\right)\right]^\frac{1}{2}}, \\
t_m(\hat{\gamma}_{1i}) & \xrightarrow{d} \frac{\tilde{u}\tilde{z}_i + \sqrt{\lambda_i} \sqrt{w'w} \tilde{r}_i}{\left[\lambda_i w'w + \tilde{z}_i^2 + \tilde{u}^2 \left(1 + \frac{\tilde{z}_i^2}{w'w}\right) + \frac{\tilde{z}_i^2 w'}{w} \right]^\frac{1}{2}}, \\
t_c(\hat{\gamma}_2) & \xrightarrow{d} \frac{\tilde{u}}{\left(1 + \frac{\tilde{u}^2 + v}{w'w}\right)^\frac{1}{2}}, \\
t_m(\hat{\gamma}_2) & \xrightarrow{d} \frac{\tilde{u}}{\left(1 + \frac{\tilde{u}^2 + v}{w'w}\right)^\frac{1}{2}},
\end{align*}
\]

where \(\lambda_i\) is a positive constant and its explicit expression is given in the Appendix.

(b) Suppose that the assumptions in Theorem 2 hold and denote the sign operator by \(\text{sgn}(\cdot)\). If \(\delta > 0\), i.e., the model is misspecified, we have

\[
\begin{align*}
t_c(\hat{\gamma}_{1i}) & \xrightarrow{d} \frac{\tilde{z}_i}{\left(1 + \frac{\tilde{z}_i^2}{w'w}\right)^\frac{1}{2}}, \\
t_m(\hat{\gamma}_{1i}) & \xrightarrow{d} N\left(0, \frac{1}{4}\right), \\
t_c(\hat{\gamma}_2) & \xrightarrow{d} \text{sgn}(s) \sqrt{w'w}, \\
t_m(\hat{\gamma}_2) & \xrightarrow{d} N(0,1).
\end{align*}
\]

Proof. See the Appendix.

**Corollary 1.**

(a) Suppose that the assumptions in part (a) of Theorem 2 hold. Then, for correctly specified models, the limiting distributions of \(t_c^2(\hat{\gamma}_{1i})\), \(t_m^2(\hat{\gamma}_{1i})\), \(t_c^2(\hat{\gamma}_2)\), and \(t_m^2(\hat{\gamma}_2)\) are stochastically dominated by \(\chi_1^2\).

(b) Suppose that the assumptions in part (b) of Theorem 2 hold. Then, for misspecified models, the limiting distributions of \(t_c^2(\hat{\gamma}_{1i})\) and \(t_m^2(\hat{\gamma}_{1i})\) are stochastically dominated by \(\chi_1^2\).

Proof. See the Appendix.
Theorem 2 and Corollary 1 illustrate the implications of using standard inference procedures (critical values from $N(0,1)$) for testing the statistical significance of the SDF parameters $\gamma$ in the presence of a useless factor. Apart from $t_m(\hat{\gamma}_2)$ in misspecified models, all the other statistics are not asymptotically distributed as standard normal random variables. For example, in misspecified models, the test statistic $t_c(\hat{\gamma}_2)$ will over-reject the null hypothesis when $N(0,1)$ is used as a reference distribution and this over-rejection increases with the number of test assets $N$. As a result, researchers will conclude erroneously (with high probability) that the factor $g_t$ is important and should be included in the model. In order to visualize the source of the over-rejection problem, Figure 1 plots the probability density function of $t_c(\hat{\gamma}_2)$ for $N - K = 7$ when the model is misspecified.

![Figure 1 about here](image)

Given the bimodal shape and a variance of 7 for the limiting distribution of $t_c(\hat{\gamma}_2)$, using the critical values from the standard normal distribution would obviously result in highly misleading inference. Importantly, part (b) of Theorem 2 shows that the $t$-statistic under potentially misspecified models, $t_m(\hat{\gamma}_2)$, retains its standard normal asymptotic distribution even when the factor is useless and Figure 1 provides a graphical illustration of this result. The reduction in the degrees of freedom from $N - K$ for the asymptotic chi-squared distribution of $t_c(\hat{\gamma}_2)^2$ to 1 for the asymptotic chi-squared distribution of $t_m(\hat{\gamma}_2)^2$ is striking.

Theorem 2 also suggests that the presence of a useless factor renders the inference on all the remaining parameters non-standard. Testing the statistical significance of the parameters on the useful factors, in both correctly specified and misspecified models, against the standard normal critical values would lead to under-rejection of the null hypothesis and conservative inference.

The main conclusion that emerges from these results is that one should use misspecification-robust $t$-statistics when testing the statistical significance of individual SDF parameters. This will ensure that the statistical decision from this test is robust to possible model misspecification and useless factors. If the model happens to be correctly specified, this will result in conservative inference but the useless factor will be removed with probability greater than $1 - \alpha$, where $\alpha$ is the size of the test. If a useless factor is not present in the model, the standard normal asymptotics for the misspecification-robust test is restored as discussed in Section 1.
Finally, it is instructive to investigate whether the presence of a useless factor affects the limiting behavior of the specification test based on the sample squared HJ-distance

$$
\hat{\delta}^2 = \hat{e}'\hat{U}^{-1}\hat{e}.
$$

(46)

In the absence of a useless factor, it is well known that under a correctly specified model (Jagannathan and Wang, 1996)

$$
T\hat{\delta}^2 \overset{d}{\to} \sum_{i=1}^{N-K} \xi_i X_i,
$$

(47)

where the $X_i$’s are independent chi-squared random variables with one degree of freedom and the $\xi_i$’s are the $N - K$ nonzero eigenvalues of

$$
\hat{S}_{1/2}U^{-1}\hat{S}_{1/2} - \hat{S}_{1/2}U^{-1}B(B'U^{-1}B)^{-1}B'U^{-1}\hat{S}_{1/2}.
$$

(48)

In practice, the specification test based on the HJ-distance is performed by comparing $T\hat{\delta}^2$ with the critical values of $\sum_{i=1}^{N-K} \hat{\xi}_i X_i$, where the $\hat{\xi}_i$’s are the nonzero eigenvalues of

$$
\hat{S}_{1/2}U^{-1}\hat{S}_{1/2} - \hat{S}_{1/2}U^{-1}\hat{B}(\hat{B}'U^{-1}\hat{B})^{-1}\hat{B}'U^{-1}\hat{S}_{1/2}.
$$

(49)

When the model is misspecified, Hansen, Heaton, and Luttmer (1995) show that the sample squared HJ-distance has a limiting normal distribution. However, in the presence of a useless factor, the above results will not hold. In the next theorem, we add to the existing literature (Kan and Zhang, 1999b) by characterizing the limiting behavior of the sample squared HJ-distance in the presence of a useless factor.

**Theorem 3.** Let $Q_1 \sim \text{Beta}\left(\frac{N-K}{2}, \frac{1}{2}\right)$ with density $f_{Q_1}(\cdot)$, $Q_2 \sim \text{Beta}\left(\frac{N-K-1}{2}, \frac{1}{2}\right)$ with density $f_{Q_2}(\cdot)$ and $c_\alpha$ be the $100(1-\alpha)$-th percentile of $\chi^2_{N-K-1}$.

(a) Suppose that the assumptions in part (a) of Theorem 2 hold. If $H_0 : \delta^2 = 0$, we have

$$
T\hat{\delta}^2 \overset{d}{\to} E[\tilde{f}(\gamma_1)^2]\chi^2_{N-K-1}
$$

(50)

and the limiting probability of rejecting $H_0 : \delta^2 = 0$ by the HJ-distance test of size $\alpha$ is

$$
\int_0^1 \left[ \chi^2_{N-K-1} > \frac{c_\alpha}{q} \right] f_{Q_1}(q) dq < \alpha.
$$

(51)
(b) Suppose that the assumptions in Theorem 1 hold. If $\delta > 0$, we have

$$\frac{\delta^2}{\chi^2_d} \to \delta^2 Q_2$$

and the limiting probability of rejecting $H_0 : \delta^2 = 0$ by the HJ-distance test of size $\alpha$ is

$$\int_0^1 P\left[\chi^2_{N-K} > \frac{c_0 q}{1-q}\right] f_{Q_2}(q) dq < 1.$$  

\textbf{Proof.} See the Appendix.

An immediate consequence of the result in Theorem 3 is that the presence of a useless factor tends to distort the inference on the specification test as well. More specifically, part (b) of Theorem 3 reveals that the HJ-distance test of correct model specification is inconsistent under the alternative.

Note that the limiting probabilities of rejection in (51) and (53) are only functions of the significance level $\alpha$ and the degree of over-identification $N - K$. Figure 2 plots these probabilities for different significance levels ($\alpha = 0.01, 0.05, \text{and } 0.1$) and $N - K$ ranging from 2 to 20.

The top panel of Figure 2 reveals that under a correctly specified model, the limiting probability of rejection of the HJ-distance test is below its nominal level when a useless factor is present. When the model is misspecified, the bottom panel of Figure 2 shows that the probability of rejection of the HJ-distance test will not approach one even in large samples. In fact, there is a nonzero probability that the HJ-distance test will favor the null of correct specification, and this probability is particularly high when $N - K$ is small. As a result, the presence of a useless factors makes it more difficult for the HJ-distance test to detect a misspecified model.

Overall, our theoretical results suggest that using the misspecification-robust $t$-test of zero risk premium would be a convenient tool for identifying if a factor is useless. While it might be desirable to develop an inference procedure on the remaining SDF parameters that is fully robust to the presence of useless factors, this does not seem to be feasible in our framework. In fact, the presence of a useless factor distorts the standard inference on the remaining SDF parameters, their associated $t$-statistics and the model specification test. We show that the presence of a useless
factor renders the remaining parameter estimates inconsistent and causes their $t$-statistics under both correct model specification and model misspecification to under-reject the null. Only after the useless factor is identified and removed using the misspecification-robust $t$-test, the validity of the inference and the consistency of the parameters are restored.

These considerations suggest that a sequential procedure based on the misspecification-robust $t$-statistics is necessary. Specifically, in the first stage, we estimate the full model and the factor with the smallest (in absolute value) $t$-statistic, for which the null of zero risk premium is not rejected at the pre-specified nominal level, is eliminated from the model. The model is then re-estimated with only the factors that survive the first stage. This procedure is repeated until either all factors are eliminated or the SDF parameter estimates on the remaining factors are found to be statistically significant at the desired nominal level when using the misspecification-robust $t$-test. The appeal of this model selection procedure is that testing the significance of the SDF parameters $\gamma$ is based on the critical values from the $N(0,1)$ distribution. The effectiveness of our proposed method in eliminating useless factors (and factors with zero risk premia) and retaining useful factors in the model is analyzed in the simulation section below.\footnote{Instead of eliminating factors with insignificant $t$-ratios one at a time, one may be tempted to drop all the factors with insignificant $t$-ratios in each stage. Unlike our proposed method, this alternative procedure can lead to the undesirable outcome of eliminating multiple useful factors when a linear combination of them is useless. In this situation, only one of these useful factors should be dropped to restore the full rank condition for the remaining factors.}

3. Monte Carlo Simulations

In this section, we undertake a Monte Carlo experiment to assess the small-sample properties of the various test statistics in models with useful and useless factors. In our simulations, we also evaluate the effectiveness of the sequential model selection procedure described above in retaining useful factors and eliminating useless factors and factors with zero risk premia.

3.1 Tests of parameter restrictions

For the analysis of the SDF parameter and specification tests, we consider three linear models: (i) a model with a constant term and a useful factor, (ii) a model with a constant term and a useless factor, and (iii) a model with a constant term, a useful factor and a useless factor. For each model,
we consider two separate cases: the case in which the model is correctly specified and the case in which the model is misspecified. The returns on the test assets and the useful factor are drawn from a multivariate normal distribution. In all simulation designs, the covariance matrix of the simulated test asset returns is set equal to the estimated covariance matrix from the 1959:2–2007:7 sample of monthly gross returns on the one-month T-bill and the 25 Fama-French size and book-to-market ranked portfolios (from Kenneth French’s website). For misspecified models, the means of the simulated returns are set equal to the sample means of the actual returns. For correctly specified models, the means of the simulated returns are set such that the asset pricing model restrictions are satisfied (i.e., the pricing errors are zero). For the simulated useful factor, we calibrate its mean and variance to the sample mean and variance of the value-weighted market excess return. The covariances between the useful factor and the returns are chosen based on the covariances estimated from the data. The useless factor is generated as a standard normal random variable independent of the returns and the useful factor. The time-series sample size is taken to be $T = 200, 600, \text{ and } 1000$. These choices of $T$ cover the range of sample sizes that are typically encountered in empirical work. We also present the limiting rejection probabilities based on our asymptotic results in Theorems 2 and 3.

In Tables 1 to 3, we report the probabilities of rejection (based on 100,000 simulations) of $H_0 : \gamma_i = \gamma_i^*$ for models (i), (ii), and (iii), respectively, where the $\gamma_i^*$’s for the constant and the useful factor are the chosen pseudo SDF parameters, and the $\gamma_i^*$ for the useless factor is set equal to zero. We present results by comparing two different $t$-statistics with the standard normal distribution, the one computed under the assumption that the model is correctly specified, $t_c(\hat{\gamma}_i)$, and the one computed under the assumption that the model is potentially misspecified, $t_m(\hat{\gamma}_i)$. For each table, Panel A reports the probabilities of rejection when the model is correctly specified and Panel B reports the probabilities of rejection when the model is misspecified.

The results in Table 1.A show that for models that are correctly specified and contain only useful factors, the standard asymptotics provides an accurate approximation of the finite-sample behavior

---

8 See Gospodinov, Kan, and Robotti (2013) for a detailed description of how the parameters are chosen in the different simulation designs.

9 The limiting rejection probabilities of $t_c$ in Table 1.B are computed based on (16) assuming that the factor and the returns are multivariate normally distributed.

---
of the $t$-tests. Since the useful factor, calibrated to the properties of the value-weighted market excess return, is closely replicated by the returns on the test assets, the differences between the $t$-tests under correctly specified models ($t_c$) and the $t$-tests under potentially misspecified models ($t_m$) are negligibly small even when the model fails to hold exactly (see Panel B).

Tables 2 and 3 present the empirical size of the $t$-tests in the presence of a useless factor. The simulation results for the $t$-tests on the parameters of the useful factor (and the constant term) confirm our theoretical findings that the null hypothesis is under-rejected when $N(0, 1)$ is used as a reference distribution. This is the case for correctly specified and misspecified models.

Similarly, the inference on the useless factor proves to be conservative when the model is correctly specified. However, when the model is misspecified, there are substantial differences between $t_c$ and $t_m$ for the useless factor. Under this scenario, we argued in Section 2 that the $t$-statistics under correct model specification have a non-normal asymptotic distribution while the misspecification-robust $t$-statistic for the parameter on the useless factor has a $N(0, 1)$ asymptotic distribution. Since the $t_c$ test on the useless factor is asymptotically distributed (up to a sign) as $\sqrt{\chi^2_{N-K}}$, it tends to over-reject severely when the critical values from $N(0, 1)$ are used and the degree of over-rejection increases with the sample size. In contrast, the $t_m$ test on the useless factor has good size properties although, for small sample sizes, it slightly under-rejects. As the sample size increases, the empirical rejection rates approach the limiting rejection probabilities (as shown in the rows for $T = \infty$) computed from the corresponding asymptotic distributions in Theorem 2.

### 3.2 Specification test

As shown in Theorem 3, the model specification test based on the HJ-distance is not immune to the useless factor problem and will be inconsistent under the alternative hypothesis of model misspecification when a useless factor is present. To illustrate the differences in the rejection probabilities for different number of assets, we also report results for the 10 size portfolio returns obtained from Kenneth French’s website.
The results in Table 4 support our theoretical findings that the probability of rejecting the null hypothesis is a function of the number of test assets and tends to be lower when $N-K$ is smaller. As the sample size increases, all empirical rejection rates approach the limiting rejection probabilities (as shown in the rows for $T = \infty$) computed from the corresponding asymptotic representations in Theorem 3.

### 3.3 Model selection procedure

Our findings suggest that the misspecification-robust $t$-test should always be used, regardless of whether a factor is deemed to be useful or useless and a model is considered to be correctly specified or misspecified. However, since the procedure based on the $t_m$ test is often conservative, some useful factors might be erroneously excluded from the model. The frequency at which this happens is evaluated in the model selection procedure presented below.

Table 5 reports the survival rates of different factors when using the sequential procedure described at the end of Section 2. In particular, we compare the survival rates from using the $t_m$ test to the survival rates from using the $t_c$ test. In our simulations, we consider a linear model with a constant term, a useful factor with $\gamma_i^* \neq 0$, a useful factor with $\gamma_i^* = 0$, and a useless factor. As in Tables 1–4, the returns and the factors are drawn from a multivariate normal distribution. The mean and variance of the useful factor with $\gamma_i^* \neq 0$ are calibrated to the mean and variance of the excess market return. The mean and variance of the useful factor with $\gamma_i^* = 0$ are calibrated to the mean and variance of the small-minus-big factor of Fama and French (1993). Finally, the useless factor is generated as a standard normal random variable, independent of the test asset returns and the useful factors. The time-series sample size is taken to be $T = 200, 600$, and $1000$. The nominal level of the sequential testing procedure is set equal to 5%.

Table 5 about here

Panel A shows that when the model is correctly specified, the procedures based on $t_c$ and $t_m$ do a similarly good job in retaining the useful factor in the model and eliminating the useless factor and the factor that does not reduce the HJ-distance. This indicates that using the $t_c$ test in the presence of a useless factor is not problematic when the underlying model holds exactly. However, as shown in Panel B, the situation drastically changes when the model is misspecified. In this case,
the procedures based on $t_c$ and $t_m$ still retain the useful factor with similarly high probability (as
the sample size gets larger), but they produce very different results when it comes to the useless
factor. For example, the procedure based on $t_c$ will retain the useless factor 48% of the time for
$T = 1000$. In contrast, the procedure based on $t_m$ will retain the useless factor only about 4.7% of
the time for $T = 1000$. It should be emphasized that the effectiveness of the proposed sequential
procedure in retaining the useful factor in the model depends on the correlation between the useful
factor and the returns on the test assets and on the magnitude of the SDF coefficient associated
with the useful factor. Our procedure will be more effective in retaining a useful factor in the
model, the higher this correlation and the larger the SDF coefficient on the useful factor.\footnote{Results for the properties of our selection procedure when the model includes factors that are only weakly
correlated with the test asset returns are available from the authors upon request. These additional simulation
results also investigate how well our proposed asymptotics characterizes the less extreme case of a factor that exhibits
a low (but nonzero) correlation with the returns.}

Some aspects of our model selection procedure deserve additional remarks. It should be stressed
that the objective of our proposed sequential test is to find the most parsimonious model with the
same HJ-distance as the full model. The fact that our method eliminates unpriced useful factors
along with useless factors is not of particular concern since these factors do not contribute to the
reduction of the HJ-distance. As a result, our model selection procedure purges the model from
useless factors (that give rise to inference problems) and unpriced factors (that do not improve the
pricing ability of the model) and restores the standard asymptotic theory for the remaining factors
in the SDF.

Finally, we consider a scenario in which a linear combination of two useful factors is useless.
Although our theoretical setup in Section 2 is not specifically designed to deal with this type of
situation, it is still interesting to examine how our sequential model selection procedure fares in
this framework. Each factor is created by adding a normally distributed error to the excess market
return. The error term in each factor has a mean of zero and a variance of 4% of the variance of the
excess market return. The two error terms are independent of each other as well as of the returns
on the test assets and the market portfolio. As in Table 5, the returns and the factors are drawn
from a multivariate normal distribution. We are interested in determining the probability that (i)
both factors survive, (ii) only one factor survives, and (iii) no factor survives using the sequential
procedure based on misspecification-robust $t$-tests. For comparison, we also report results of the
sequential procedure based on $t$-tests under correct model specification. The nominal level of the
sequential testing procedures is set equal to 5%. Ideally, in this framework, only one factor should survive the testing procedures described above.

Panel A of Table 6 shows that when the model is correctly specified, the procedures based on $t_c$ and $t_m$ do a similarly good job in retaining only one factor in the model. For example, for $T = 1000$, the probability that only one factor survives is either 93% or 95% depending on whether we use $t_c$ or $t_m$. For this sample size, the probabilities that both factors survive and no factor survives are very low and similar across procedures. However, when the model is misspecified (see Panel B), the procedures based on $t_c$ and $t_m$ deliver very different results. For $T = 1000$, the probability that both factors survive the model selection procedure based on $t_c$ is about 46% while the probability that both factors survive the model selection procedure based on $t_m$ is about 4%. This difference in probabilities becomes larger as the sample size is allowed to grow. Importantly, the probabilities that only one factor survives are markedly different across procedures. For example, when $T = 1000$, the probability that only one factor survives is about 92% when using $t$-tests under misspecified models while it is only about 52% when using $t$-tests under correctly specified models. In summary, our selection procedure based on $t$-tests that are robust to misspecification continues to perform reasonably well even when no single factor is useless but a linear combination of them is.

4. Empirical Analysis

Our theoretical and simulation results point out some serious pitfalls in the empirical analysis of asset pricing models with non-traded factors. In the following, we use monthly data to demonstrate the relevance of our theoretical results. To show that our findings are not specific to the SDFs considered in the main empirical application, we also use a quarterly dataset to analyze the performance of additional models with non-traded factors.\textsuperscript{11}

\textsuperscript{11}Some additional empirical results are provided in an online appendix available on the authors' websites.
4.1 Main application

First, we describe the data used in the empirical analysis and outline the different specifications of the asset pricing models considered. Then, we present our results.

4.1.1 Data and asset pricing models

As in the Monte Carlo simulations, the test asset returns are the monthly gross returns on the one-month T-bill and the value-weighted 25 Fama-French size and book-to-market ranked portfolios from February 1959 until July 2007. We analyze seven asset pricing models starting with the simple static CAPM. The SDF specification for this model is

\[ y_t^{CAPM}(\gamma) = \gamma_0 + \gamma_1vw_t, \]

where \( vw \) is the excess return (in excess of the one-month T-bill rate) on the value-weighted stock market index (NYSE-AMEX-NASDAQ) from Kenneth French’s website. The CAPM performed well in the early tests, e.g., Fama and MacBeth (1973), but has fared poorly since.

One extension that has performed better is our second model, the conditional CAPM (C-LAB) of Jagannathan and Wang (1996). This model incorporates measures of the return on human capital as well as the change in financial wealth and allows the conditional SDF coefficients to vary with a state variable, \( prem \), the lagged yield spread between Baa and Aaa rated corporate bonds from the Board of Governors of the Federal Reserve System.\(^{12}\) The SDF specification is

\[ y_t^{C-LAB}(\gamma) = \gamma_0 + \gamma_1vw_t + \gamma_2prem_{t-1} + \gamma_3lab_t, \]

where \( lab \) is the growth rate in per capita labor income, \( L \), defined as the difference between total personal income and dividend payments, divided by the total population (from the Bureau of Economic Analysis). Following Jagannathan and Wang (1996), we use a two-month moving average to construct the growth rate \( lab_t = (L_t-1 + L_t-2)/(L_t-2 + L_t-3) - 1 \), for the purpose of minimizing the influence of measurement error.

Our third model (FF3) extends the CAPM by including two empirically-motivated factors. This is the Fama-French (1993) three-factor model with

\[ y_t^{FF3}(\gamma) = \gamma_0 + \gamma_1vw_t + \gamma_2smb_t + \gamma_3hml_t, \]

\(^{12}\)All bond yield data are from this source unless noted otherwise.
where \( smb \) is the return difference between portfolios of stocks with small and large market capitalizations, and \( hml \) is the return difference between portfolios of stocks with high and low book-to-market ratios ("value" and "growth" stocks, respectively) from Kenneth French’s website.

The fourth model (ICAPM) is an empirical implementation of Merton’s (1973) intertemporal extension of the CAPM based on Campbell (1996), who argues that innovations in state variables that forecast future investment opportunities should serve as factors. The five-factor specification proposed by Petkova (2006) is

\[
y_t^{ICAPM}(\gamma) = \gamma_0 + \gamma_1 vw_t + \gamma_2 term_t + \gamma_3 def_t + \gamma_4 div_t + \gamma_5 rf_t,
\]

where \( term \) is the difference between the yields of ten-year and one-year government bonds, \( def \) is the difference between the yields of long-term corporate Baa bonds and long-term government bonds (from Ibbotson Associates), \( div \) is the dividend yield on the Center for Research in Security Prices (CRSP) value-weighted stock market portfolio, and \( rf \) is the one-month T-bill yield (from CRSP, Fama Risk Free Rates). The actual factors for \( term, def, div, \) and \( rf \) are their innovations from a VAR(1) system of seven state variables that also includes \( vw, smb, \) and \( hml \).

Next, we consider consumption-based models. Our fifth model (CCAPM) is the unconditional consumption model with

\[
y_t^{CCAPM}(\gamma) = \gamma_0 + \gamma_1 c_t,
\]

where \( c \) is the growth rate in real per capita total consumption (seasonally adjusted at annual rates) from the Bureau of Economic Analysis. This model has generally not performed well empirically. Therefore, we also examine two other consumption models that have yielded more encouraging results.

One such model (CC-CAY) is a conditional version of the CCAPM due to Lettau and Ludvigson (2001). The relation is

\[
y_t^{CC-CAY}(\gamma) = \gamma_0 + \gamma_1 c_t + \gamma_2 cay_{t-1} + \gamma_3 c_t \cdot cay_{t-1},
\]

where \( cay \), the conditioning variable, is a consumption-aggregate wealth ratio.\(^{13}\) This specification is obtained by scaling the constant term and the \( c \) factor of a linearized consumption CAPM by a

\(^{13}\)Following Jørgensen and Attanasio (2003), we linearly interpolate the quarterly values of \( cay \) to permit analysis at the monthly frequency.
constant and $\text{cay}$. Scaling factors by instruments is one popular way of allowing factor risk premia to vary over time. See Cochrane (1996), among others.

The last model (D-CCAPM), due to Yogo (2006), highlights the cyclical role of durable consumption in asset pricing. The specification is

$$y_{t-CCAPM}^{D}(\gamma) = \gamma_0 + \gamma_1 w_t + \gamma_2 c_{nd,t} + \gamma_3 c_{d,t},$$

where $c_{nd}$ ($c_d$) is the growth rate in real per capita nondurable (durable) consumption (seasonally adjusted at annual rates) from the Bureau of Economic Analysis.

### 4.1.2 Results

Before presenting the estimation results for the SDF parameters, we first investigate whether the various risk factors are correlated with the asset returns and whether the seven models described above are properly identified. As mentioned in the theoretical section of the paper, the presence of a useless factor leads to a violation of the crucial identification condition that the $N \times K$ matrix $B = E[x_t \tilde{f}_t]$ is of full column rank. Therefore, it is of interest to test if $B$ is of (reduced) rank $K - 1$. Since $\hat{B}$ is not invariant to rescaling of the data, we first perform a normalization on $\hat{B}$. Define $\hat{\Theta} = \hat{U}^{-\frac{1}{2}} \hat{S} \hat{V}^{-\frac{1}{2}}$ and its corresponding population counterpart $\Theta = U^{-\frac{1}{2}} BS^{-\frac{1}{2}}$, where $\hat{S}_j = \frac{1}{T} \sum_{t=1}^{T} \hat{f}_t \hat{f}_t'$ and $S_j = E[\hat{f}_t \hat{f}_t']$. Note that $\hat{\Theta}$ is invariant to rescaling of the data and rank($\Theta$) = rank($B$).

Let $\hat{\Pi}$ be a consistent estimator of the asymptotic covariance matrix of $\sqrt{T} \text{vec}(\hat{\Theta} - \Theta)$, where $\text{vec}(\cdot)$ is the vec operator. Following Ratsimalahelo (2002) and Kleibergen and Paap (2006), we perform a singular value decomposition on $\hat{\Theta}$ such that $\hat{\Theta} = \tilde{U} \tilde{S} \tilde{V}'$, where $\tilde{U}' \tilde{U} = I_N$, $\tilde{V}' \tilde{V} = I_K$, and $\tilde{S}$ is an $N \times K$ matrix with the singular values of $\hat{\Theta}$ in decreasing order on its diagonal. Let $\tilde{U}_2$ be the last $N - K + 1$ columns of $\tilde{U}$, $\tilde{V}_2$ be the last column of $\tilde{V}$, and

$$\hat{\Pi} = (\tilde{V}_2' \otimes \tilde{U}_2') \hat{\Pi}(\tilde{V}_2 \otimes \tilde{U}_2).$$

Then, a test of $H_0 : \text{rank}(\Theta) = \text{rank}(B) = K - 1$ is given by

$$W^* = T \tilde{s}_K^2 \hat{\Pi}^{11} \overset{d}{\rightarrow} \chi^2_{N-K+1},$$

where $\tilde{s}_K$ is the smallest singular value of $\hat{\Theta}$, and $\hat{\Pi}^{11}$ is the (1,1) element of $\hat{\Pi}^{-1}$.
Table 7.A reports the rank restriction test ($W^*$) and its $p$-value ($p$-val) of the null that $E[x_t(1, f_{ii})]$ has a column rank of one. The panel shows that we cannot reject the null of a column rank of one at the 5% significance level for eight out of 14 macroeconomic and financial factors. This finding suggests that several risk factors can be reasonably considered as useless and that our asymptotic results on useless factors are of practical importance. Panel B further shows that only CAPM and FF3 convincingly pass the test of full rank condition. This is consistent with the fact that $vw$, $smb$ and $hml$ are highly correlated with the returns on the test assets while most factors in the other models are not. Panel B also shows that only ICAPM passes the HJ-distance specification test at conventional levels of significance. Since the HJ-distance test has been shown to substantially over-reject under the null in realistic simulation settings with many test assets, we also consider an alternative test of $H_0 : \lambda = U^{-1}e = 0_N$ (which is equivalent to the test of $H_0 : \delta = 0$). Gospodinov, Kan, and Robotti (2013) show that this alternative Lagrange multiplier ($LM$) test has excellent size properties. The results in Panel B indicate that one would reach the same conclusions using the $LM$ and HJ-distance tests. Therefore, the model rejections documented in Table 7.B do not seem to be driven by the finite-sample properties of the HJ-distance test. Overall, these empirical findings suggest that valid inference should account for the fact that models are potentially misspecified and poorly identified.

Although the rank restriction test serves as a useful pre-test for possible identification problems, this test does not allow us to unambiguously identify which factor contributes to the identification failure of the model. In addition, this test does not address the question of which risk factors are important in explaining the cross-sectional differences in asset returns. Our misspecification-robust test of $H_0 : \gamma_i = 0$ proves to be of critical importance in (i) providing the direction of the identification failure and (ii) allowing us to determine whether a given risk factor is priced. Panels C and D of Table 7 present the $t$-tests under correct model specification and potential model misspecification as well as the model selection procedure described at the end of Section 2. Using $t$-tests under correct specification, the results in Panel C suggest that the factors $lab$ in C-LAB, $term$ in ICAPM, $c$ in CCAPM and CC-CAY, $cay$ in CC-CAY, and $c_{rod}$ in D-CCAPM survive the sequential test procedure at the 5% significance level. However, given the violation of the full rank condition for these models, the standard normal distribution is not the appropriate reference.
distribution in this case. Since some of these factors are very weakly correlated with the returns on the test assets and effectively behave as useless factors, they tend to be included in the model much more often than they should (see the second to last column in Panel B of Table 5). Therefore, the model selection procedure needs to be implemented using misspecification-robust \( t \)-tests.

Panel D shows that, from all of the non-traded factors listed above, only \( \text{term} \) in ICAPM and \( c_{nd} \) in D-CCAPM survive the sequential procedure based on misspecification-robust \( t \)-tests at the 5\% significance level. Finally, for traded factors, we find strong evidence of pricing for the \( \text{vw} \) factor in CAPM and the \( \text{vw} \) and \( \text{hml} \) factors in FF3 in both Panels C and D.

### 4.2 Additional empirical evidence

In this subsection, we analyze the performance of some prominent asset pricing models using quarterly data. The test asset returns are the quarterly gross returns on the one-month T-bill and the value-weighted 25 Fama-French size and book-to-market ranked portfolios from 1952 Q2 until 2007 Q4. In addition to CAPM, FF3, CCAPM, CC-CAY and D-CCAPM, we consider the following asset pricing specifications: (i) the conditional CAPM (C-ML) of Santos and Veronesi (2006) with \( \text{vw} \) and \( \text{vw} \cdot \text{ml} \) scaled by the labor income-consumption ratio (\( \text{ml} \)) as risk factors; (ii) a version of the conditional consumption CAPM (CC-MY) proposed by Lustig and Van Nieuwerburgh (2005) with the housing collateral ratio (\( \text{my} \)), \( c \), and the interaction term \( \text{my} \cdot c \) as risk factors; and (iii) the sector investment model (SIM) of Li, Vassalou, and Xing (2006) with the log investment growth rates for households (\( i_h \)), non-financial corporations (\( i_c \)), and non-corporate sector (\( i_{nc} \)) as risk factors. These three additional models with non-traded factors have yielded encouraging results in cross-sectional asset pricing.

The empirical results for quarterly data are reported in Table 8, with Panel A showing that for all factors except for \( \text{vw} \), \( \text{vw} \cdot \text{ml} \), \( \text{smb} \), \( \text{hml} \), and \( c_{d} \), we cannot reject the null that \( E[x_t(1, f_{it})] \) has a column rank of one at the 5\% significance level. In addition, the results in Panel B indicate that we cannot reject the null of reduced rank for all models except for CAPM and FF3 and that all models except for SIM are rejected by the HJ-distance specification test.\(^1\) This clearly points to the need of statistical procedures that are robust to model misspecification and weak identification.

---

\(^1\)Using the \( LM \) test and a 5\% significance level, we can reject the null of correct specification for all models.
Panel D of Table 8 shows that all the non-traded factors except for \( cay \) and \( vw \cdot ml \) do not survive the model selection procedure based on the misspecification-robust \( t \)-test. This stands in sharp contrast to the results in Panel C of Table 8 where the \( t \)-test under correctly specified models is employed. However, our theoretical and simulation analyses clearly showed that relying on the \( t \)-test under correct specification is grossly inappropriate when the underlying model is misspecified and the factors are very weakly correlated with the returns on the test assets. As one example, consider CC-MY. In the final stage of the model selection procedures in Panels C and D, the \( t \)-statistic on the risk premium estimator for the \( my \) factor drops from 2.36 to 1.44 when misspecification and weak identification are taken into account.

Finally, as in the monthly case, the two traded factors \( vw \) and \( hml \) appear to be priced in the cross-section of asset returns at the 5% significance level. Taken together, these results serve as a warning signal to researchers that are interested in estimating and analyzing SDF parameters on non-traded risk factors.

5. Conclusion

It is well known that asset returns are, at best, only weakly correlated with many macroeconomic factors. Nonetheless, researchers in finance have typically relied on inference methods that are not robust to weak identification and model misspecification when evaluating the incremental pricing ability of these factors. Our paper demonstrates that when a model is misspecified, the standard \( t \)-test of statistical significance will lead us to erroneously conclude, with high probability, that a useless factor is relevant and should be included in the model. Importantly, we show that the \( t \)-test of statistical significance will be valid only if it is computed using misspecification-robust standard errors. Furthermore, we argue that the presence of a useless factor affects the inference on the remaining model parameters and the test of correct specification. In particular, when a useless factor is present in the model, the limiting distributions of the \( t \)-statistics for the useful factors are non-standard and the HJ-distance specification test is inconsistent.

In order to overcome these problems, we propose an easy-to-implement sequential model selection procedure based on misspecification-robust \( t \)-tests that restores the standard inference on
the parameters of interest. We show via simulations that the proposed procedure is effective in eliminating useless factors as well as factors that do not improve the pricing ability of the model.

Finally, we employ our methodology to investigate the empirical performance of several prominent asset pricing models with traded and non-traded factors. While the market factor and the book-to-market factor of Fama and French (1993) are often found to be priced, the statistical evidence on the pricing ability of many non-traded factors is rather weak when using the model selection procedure based on misspecification-robust $t$-tests.
Appendix: Preliminary Lemma and Proofs of Main Results

A.1 Preliminary Lemma

Lemma A.1. Let

\[ x_t = BS_f^{-1} \tilde{f}_t + \epsilon_t, \]  

(A.1)

where \( B = E[x_t \tilde{f}_t], \) \( S_f = E[\tilde{f}_t \tilde{f}_t'] \) and \( E[\epsilon_t | \tilde{f}_t] = 0_N. \) Suppose \( \text{Cov}[\epsilon_t \epsilon'_t, (\tilde{f}_t \gamma_1^*)^2] = 0_{N \times N} \) (a sufficient condition for this to hold is \( E[\epsilon_t \epsilon'_t | \tilde{f}_t] = \Sigma, \) i.e., conditional homoskedasticity). When the model is correctly specified, we have

\[ S = E[(x_t \tilde{f}_t \gamma_1^* - q)(x_t \tilde{f}_t \gamma_1^* - q)'] = E[(\tilde{f}_t \gamma_1^*)^2] U + BCB', \]

(A.2)

where \( U = E[x_t x'_t] \) and \( C \) is a symmetric \( K \times K \) matrix.

Proof of Lemma A.1. Under a correctly specified model, we have \( q = B \gamma_1^*. \) It follows that

\[ S = E[x_t x'_t (\tilde{f}_t \gamma_1^*)^2] - qq' = E[x_t x'_t (\tilde{f}_t \gamma_1^*)^2] - B \gamma_1^* \gamma_1^* B'. \]

(A.3)

For the first term, we have

\[ E[x_t x'_t (\tilde{f}_t \gamma_1^*)^2] = E[x_t x'_t] E[(\tilde{f}_t \gamma_1^*)^2] + \text{Cov}[x_t x'_t, (\tilde{f}_t \gamma_1^*)^2] \]

\[ = E[(\tilde{f}_t \gamma_1^*)^2] U + \text{Cov}[B S_f^{-1} \tilde{f}_t \tilde{f}_t S_f^{-1} B' + \epsilon_t \epsilon'_t, (\tilde{f}_t \gamma_1^*)^2] \]

\[ = E[(\tilde{f}_t \gamma_1^*)^2] U + B S_f^{-1} \text{Cov}[\tilde{f}_t \tilde{f}_t', (\tilde{f}_t \gamma_1^*)^2] S_f^{-1} B', \]

(A.4)

where the last equality follows from the assumption that \( \text{Cov}[\epsilon_t \epsilon'_t, (\tilde{f}_t \gamma_1^*)^2] = 0_{N \times N}. \) Therefore, we have

\[ S = E[(\tilde{f}_t \gamma_1^*)^2] U + BCB', \]

(A.5)

where

\[ C = S_f^{-1} \text{Cov}[\tilde{f}_t \tilde{f}_t', (\tilde{f}_t \gamma_1^*)^2] S_f^{-1} - \gamma_1^* \gamma_1'^*. \]

(A.6)

This completes the proof.

A.2 Proofs of Theorems and Corollary 1

Proof of Theorem 1.
part (a): We start with the limiting distribution of $\sqrt{T}(\hat{\gamma}_1 - \gamma_1^*)$. Under the assumptions in Theorem 1, we have

$$\sqrt{T} \hat{U}^{-\frac{1}{2}}d \overset{d}{\to} z \sim N(0_N, I_N) \quad (A.7)$$

and

$$-\sqrt{T} \hat{U}^{-\frac{1}{2}}(B\hat{\gamma}_1^* - q) \overset{d}{\to} y \sim N(0_N, V_y), \quad (A.8)$$

where $V_y = E[m_tm_t']$ is the covariance matrix of $y$, and

$$m_t = U^{-\frac{1}{2}}(x_t\hat{f}_t^*\gamma_1^* - q) = U^{-\frac{1}{2}}e_t(\gamma_1^*). \quad (A.9)$$

Therefore, we have $V_y = U^{-\frac{1}{2}}SU^{-\frac{1}{2}}$ for correctly specified models. In addition, $y$ and $z$ are independent of each other. Using $y$ and $z$, we can write (24) as

$$\sqrt{T}(\hat{\gamma}_1 - \gamma_1^*) = (\hat{B}'\hat{U}^{-\frac{1}{2}}[I_N - \hat{U}^{-\frac{1}{2}}d(d'\hat{U}^{-1}d)^{-1}\hat{U}^{-\frac{1}{2}}]U^{-\frac{1}{2}}B)^{-1}$$

$$\times \hat{B}'\hat{U}^{-\frac{1}{2}}[I_N - \hat{U}^{-\frac{1}{2}}d(d'\hat{U}^{-1}d)^{-1}\hat{U}^{-\frac{1}{2}}]\sqrt{T}\hat{U}^{-\frac{1}{2}}(q - \hat{B}\gamma_1^*) \overset{d}{=} (\hat{B}'I_N - z(z'z)^{-1}z')\hat{B}^{-1}y \quad (A.10)$$

Let $w = P'z \sim N(0_{N-K}, I_{N-K})$, $u = P'y \sim N(0_{N-K}, V_u)$ with $V_u = P'U^{-\frac{1}{2}}SU^{-\frac{1}{2}}P$, $r = (\hat{B}'\hat{B})^{-\frac{1}{2}}\hat{B}'y \sim N(0_K, V_r)$ with $V_r = (\hat{B}'\hat{B})^{-\frac{1}{2}}\hat{B}'U^{-\frac{1}{2}}SU^{-\frac{1}{2}}(\hat{B}'\hat{B})^{-\frac{1}{2}}$. Making use of the identity

$$(\hat{B}'I_N - z(z'z)^{-1}z')\hat{B}^{-1} = (\hat{B}'\hat{B})^{-1} + (\hat{B}'\hat{B})^{-1}\hat{B}'zz'PP'\hat{B}(\hat{B}'\hat{B})^{-1} \quad (A.11)$$

and $z'z = z'\hat{B}(\hat{B}'\hat{B})^{-1}\hat{B}'z + w'w$, we obtain

$$\sqrt{T}(\hat{\gamma}_1 - \gamma_1^*) \overset{d}{=} (\hat{B}'\hat{B})^{-\frac{1}{2}}\left[ -\frac{w'u}{w'w}(\hat{B}'\hat{B})^{-\frac{1}{2}}\hat{B}'z + r \right]. \quad (A.12)$$

For the derivation of the limiting distribution of $\hat{\gamma}_2$, we define $M = I_N - U^{-\frac{1}{2}}B(B'U^{-1}B)^{-1}B'U^{-\frac{1}{2}}$ and $\bar{M} = I_N - \hat{U}^{-\frac{1}{2}}\hat{B}(\hat{B}'\hat{U}^{-1}\hat{B})^{-1}\hat{B}'\hat{U}^{-\frac{1}{2}}$. Using that $\bar{M}\hat{U}^{-\frac{1}{2}}\hat{B} = 0_{N\times K}$, we obtain

$$\sqrt{T}\bar{M}\hat{U}^{-\frac{1}{2}}q = \sqrt{T}\bar{M}\hat{U}^{-\frac{1}{2}}(q - \hat{B}\gamma_1^*) \overset{d}{=} My, \quad (A.13)$$

and we can rewrite $\hat{\gamma}_2$ as

$$\hat{\gamma}_2 = \frac{(\sqrt{T}\hat{U}^{-\frac{1}{2}}d)'(\sqrt{T}\bar{M}\hat{U}^{-\frac{1}{2}}(B - \hat{B})\gamma_1^*)}{(\sqrt{T}\hat{U}^{-\frac{1}{2}}d)'\bar{M}(\sqrt{T}\hat{U}^{-\frac{1}{2}}d)}. \quad (A.14)$$
Then, from (A.7), (A.8) and \( \hat{M} \overset{p}{\to} M = PP' \), we get
\[
\hat{\gamma}_2 \overset{d}{\to} \frac{z'My}{z'Mz} = \frac{(P'z)'(P'y)}{(P'z')(P'z)} = \frac{w'u}{w'w}. \tag{A.15}
\]
This completes the proof of part (a) of Theorem 1.

**part (b):** Using the fact that \( \hat{U}^{-\frac{1}{2}} \hat{B} \overset{a.s.}{\to} \hat{B} \) and \( \sqrt{T} \hat{U}^{-\frac{1}{2}} \hat{d} \overset{d}{\to} z \), we can obtain the limiting distribution of \( \hat{\gamma}_1 \) in (24) as
\[
\hat{\gamma}_1 \overset{d}{\to} (\hat{B}'[I_N - z(z'z)^{-1}z']\hat{B})^{-1} \hat{B}'[I_N - z(z'z)^{-1}z']\hat{q} \tag{A.16}
\]
Using (A.11) and the fact that \( \gamma_1^* = (\hat{B}'\hat{B})^{-1} \hat{B}'\hat{q} \), we obtain
\[
\gamma_1 - \gamma_1^* \overset{d}{\to} \left[ (\hat{B}'\hat{B})^{-1} + \frac{(\hat{B}'\hat{B})^{-1} \hat{B}'z z' \hat{B}'(\hat{B}'\hat{B})^{-1}}{w'w} \right] \left( \hat{B}'\hat{q} - \frac{\hat{B}'z z' \hat{q}}{z'z} \right) - (\hat{B}'\hat{B})^{-1} \hat{B}'\hat{q}
\]
\[
= - (\hat{B}'\hat{B})^{-1} \hat{B}'z \frac{z' \hat{q}}{z'z} + (\hat{B}'\hat{B})^{-1} \hat{B}'z \frac{z' \hat{B}(\hat{B}'\hat{B})^{-1} \hat{B}' \hat{q}}{w'w} - (\hat{B}'\hat{B})^{-1} \hat{B}'z \frac{z' \hat{q} z' z' \hat{B}(\hat{B}'\hat{B})^{-1} \hat{B}'z}{w'w}
\]
\[
= - (\hat{B}'\hat{B})^{-1} \hat{B}'z \frac{z' \hat{q}}{w'w} + (\hat{B}'\hat{B})^{-1} \hat{B}'z \frac{z' \hat{B}(\hat{B}'\hat{B})^{-1} \hat{B}' \hat{q}}{w'w}
\]
\[
= - \frac{z'M\hat{q}}{w'w} (\hat{B}'\hat{B})^{-1} \hat{B}'z
\]
\[
= - \frac{\delta s}{w'w} (\hat{B}'\hat{B})^{-1} \hat{B}'z, \tag{A.17}
\]
and the last equality follows because \( \delta^2 = \hat{q}' PP' \hat{q} \) and \( s = \hat{q}' PP' z / (\hat{q}' PP' \hat{q})^{\frac{1}{2}} \).

For the limiting distribution of \( \hat{\gamma}_2 \), we have
\[
T^{-\frac{1}{2}} \hat{\gamma}_2 = \frac{\sqrt{T} \hat{d} \hat{U}^{-\frac{1}{2}} \hat{M} \hat{U}^{-\frac{1}{2}} q}{\sqrt{T} \hat{d} \hat{U}^{-\frac{1}{2}} \hat{M} (\sqrt{T} \hat{U}^{-\frac{1}{2}} \hat{d})} \overset{d}{\to} \frac{z'M\hat{q}}{z'Mz} = \frac{\delta s}{w'w}. \tag{A.18}
\]
This completes the proof of part (b) of Theorem 1.

**Proof of Theorem 2.**

**part (a):** Using Lemma A.1, we have
\[
S = E[(\hat{f}_1^* \hat{\gamma}_1^*)^2] U + BCB'
\tag{A.19}
\]
under the conditional homoskedasticity assumption. It follows that
\[
V_u = P'U^{-\frac{1}{2}} S U^{-\frac{1}{2}} P = E[(\hat{f}_1^* \hat{\gamma}_1^*)^2] I_{N-K}, \tag{A.20}
\]
\[
V_r = (\hat{B}'\hat{B})^{-\frac{1}{2}} \hat{B}' U^{-\frac{1}{2}} S U^{-\frac{1}{2}} \hat{B}' (\hat{B}'\hat{B})^{-\frac{1}{2}} = E[(\hat{f}_1^* \hat{\gamma}_1^*)^2] I_K + (\hat{B}'\hat{B})^{-\frac{1}{2}} C (\hat{B}'\hat{B})^{\frac{1}{2}}, \tag{A.21}
\]
\[
\text{Cov}[u,r'] = P'U^{-\frac{1}{2}} S U^{-\frac{1}{2}} \tilde{B} (\hat{B}'\hat{B})^{-\frac{1}{2}} = 0_{(N-K)\times K}. \tag{A.22}
\]
Let $\tilde{u} = w'u/(w'V_u w)^{1/2} = E[(\tilde{r}'\gamma_1^*)^2]^{-1/2}w'u/(w'w)^{1/2}$. It is easy to show that $\tilde{u} \sim N(0, 1)$ and it is independent of $w, z$ and $r$. Using $\tilde{u}$, we can simplify the limiting distribution of $\sqrt{T}(\hat{\gamma}_1 - \gamma_1^*)$ in (A.12) to

$$\sqrt{T}(\hat{\gamma}_1 - \gamma_1^*) \xrightarrow{d} -E[(\tilde{r}'\gamma_1^*)^2]^{1/2}\frac{\tilde{u}}{(w'w)^{1/2}}(\tilde{B}'\tilde{B})^{-1}\tilde{B}'z + (\tilde{B}'\tilde{B})^{-1/2}r.$$  

(A.23)

The estimated covariance matrix of $\hat{\gamma}$ for a potentially misspecified model is given by

$$\hat{V}_m(\hat{\gamma}) = \frac{1}{T^2} \sum_{t=1}^{T} \hat{h}_t\hat{h}'_t,$$

(A.24)

where

$$\hat{h}_t = (\hat{D}'\hat{U}^{-1}\hat{D})^{-1}\hat{D}'\hat{U}^{-1}\hat{e}_t + (\hat{D}'\hat{U}^{-1}\hat{D})^{-1}([\hat{r}'_t, g_t'] - \hat{D}'\hat{U}^{-1}x_t)\hat{u}_t,$$

(A.25)

and $\hat{u}_t = \hat{e}'\hat{U}^{-1}x_t$. In order to derive the limiting distribution of $\hat{h}_t$, we need to obtain the limiting representations of $(\hat{D}'\hat{U}^{-1}\hat{D})^{-1}$, $(\hat{D}'\hat{U}^{-1}\hat{D})^{-1}\hat{D}'\hat{U}^{-1}$, and $\hat{u}_t$.

It is straightforward to show that

$$\hat{D}'\hat{U}^{-1} = \begin{bmatrix} \hat{B}'U^{-\frac{1}{2}} + O_p(T^{-\frac{1}{2}}) \\ \frac{1}{\sqrt{T}}z'U^{-\frac{1}{2}} + O_p(T^{-1}) \end{bmatrix},$$

(A.26)

$$\hat{D}'\hat{U}^{-1}\hat{D} = \begin{bmatrix} \hat{B}'\hat{B} + O_p(T^{-\frac{1}{2}}) & \frac{1}{\sqrt{T}}\hat{B}'z + O_p(T^{-1}) \\ \frac{1}{\sqrt{T}}z'\hat{B} + O_p(T^{-1}) & z'z + O_p(T^{-\frac{1}{2}}) \end{bmatrix}.$$  

(A.27)

Then, using the partitioned matrix inverse formula, we have

$$(\hat{D}'\hat{U}^{-1}\hat{D})^{-1} = \begin{bmatrix} H + O_p(T^{-\frac{1}{2}}) & -\sqrt{T}(\hat{B}'\hat{B})^{-1}\hat{B}'z + O_p(1) \\ -\sqrt{T}z'(\hat{B}'\hat{B})^{-1} + O_p(1) & \frac{T}{w'w} + O_p(T^{\frac{1}{2}}) \end{bmatrix},$$

(A.28)

where

$$H = (\hat{B}'[I_N - z(z'z)^{-1}z']\hat{B})^{-1} = (\hat{B}'\hat{B})^{-1} + (\hat{B}'\hat{B})^{-1}\hat{B}'z(\hat{B}'\hat{B})^{-1}. $$

(A.29)

After simplification, we obtain

$$(\hat{D}'\hat{U}^{-1}\hat{D})^{-1}\hat{D}'\hat{U}^{-1} = \begin{bmatrix} (\hat{B}'\hat{B})^{-1}\hat{B}'U^{-\frac{1}{2}} - (\hat{B}'\hat{B})^{-1}\hat{B}'z(\hat{B}'\hat{B})^{-1} \frac{w'w}{w'w} + O_p(T^{-\frac{1}{2}}) \\ \sqrt{T}\frac{w'p'u}{w'w} + O_p(1) \end{bmatrix}. $$

(A.30)

With the above expressions, we now derive the limiting distribution of $\hat{u}_t$. Note that the vector of sample pricing errors is given by

$$\hat{e} = \hat{D}\hat{\gamma} - q = \hat{D}(\hat{D}'\hat{U}^{-1}\hat{D})^{-1}\hat{D}'\hat{U}^{-1}q - q.$$  

(A.31)
Using (A.13), (A.15), and the identity
\[ I_N - \hat{U}^{-\frac{1}{2}} \hat{D}(\hat{D}'\hat{U}^{-1}\hat{D})^{-1}\hat{D}'\hat{U}^{-\frac{1}{2}} = M - \hat{M}\hat{U}^{-\frac{1}{2}}d(\hat{D}'\hat{D})^{-1}\hat{D}'\hat{U}^{-\frac{1}{2}}M, \]  
(A.32)
we can obtain the limiting distribution of \(- \sqrt{T}\hat{U}^{-\frac{1}{2}}\hat{e}\) as
\[ - \sqrt{T}\hat{U}^{-\frac{1}{2}}\hat{e} = \sqrt{T}\hat{M}\hat{U}^{-\frac{1}{2}}q - \sqrt{T}\hat{M}\hat{U}^{-\frac{1}{2}}d\hat{\gamma}_2 \xrightarrow{d} My - Mz \frac{w'u}{w'w} = P \left( I_{N-K} - \frac{ww'}{w'w} \right) u, \]  
(A.33)
and we have
\[ \sqrt{T}\hat{u}_t \xrightarrow{d} -u' \left( I_{N-K} - \frac{ww'}{w'w} \right) P'U^{-\frac{1}{2}}x_t. \]  
(A.34)
Using (A.28), (A.30), (A.34), and the fact that
\[ \hat{e}_t = x_t(\hat{f}'_1\hat{\gamma}_1 + \hat{\gamma}_2g_t) - q = x_t\hat{f}'_1\gamma_1 - q + \frac{w'u}{w'w} + O_p(T^{-\frac{1}{2}}) \]  
(A.35)
under a correctly specified model, we can write the limiting distribution of \(\hat{h}_t = [\hat{h}'_{1t}, \hat{h}_{2t}]\)', where \(\hat{h}_{1t}\) denotes the first \(K\) elements of \(\hat{h}_t\), as
\[ \hat{h}'_{1t} \xrightarrow{d} \left[ (\hat{B}'\hat{B})^{-1}\hat{B}'U^{-\frac{1}{2}} - \frac{(\hat{B}'\hat{B})^{-1}\hat{B}'zwP'U^{-\frac{1}{2}}}{w'w} \right] \left( x_t\hat{f}'_1\gamma_1 - q + x_tg_t \frac{w'u}{w'w} \right) \]
\[ + \frac{(\hat{B}'\hat{B})^{-1}\hat{B}'z}{w'w} u' \left( I_{N-K} - \frac{ww'}{w'w} \right) P'U^{-\frac{1}{2}}x_tg_t, \]  
(A.36)
\[ \frac{\hat{h}_{2t}}{\sqrt{T}} \xrightarrow{d} \frac{1}{w'w} \frac{w'P'U^{-\frac{1}{2}}}{w'w} \left( x_t\hat{f}'_1\gamma_1 - q + x_tg_t \frac{w'u}{w'w} \right) - \frac{1}{w'w} u' \left( I_{N-K} - \frac{ww'}{w'w} \right) P'U^{-\frac{1}{2}}x_tg_t. \]  
(A.37)
Under the conditional homoskedasticity assumption, we have
\[ \frac{1}{T} \sum_{t=1}^{T} (x_t\hat{f}'_1\gamma_1 - q)(x_t\hat{f}'_1\gamma_1 - q)' \xrightarrow{a.s.} S = E[(\hat{f}'_1\gamma_1)^2]U + BCB'. \]  
(A.38)
Together with the fact that
\[ \frac{1}{T} \sum_{t=1}^{T} x_tx'_t g_t^2 \xrightarrow{a.s.} E[x_tx'_t g_t^2] = E[x_tx'_t]E[g_t^2] = U, \]  
(A.39)
we can show that the estimated misspecification-robust covariance matrix of \(\hat{\gamma}_1\) has a limiting distribution of
\[ TV_m(\hat{\gamma}_1) = \frac{1}{T} \sum_{t=1}^{T} \hat{h}'_{1t}\hat{h}_{1t}' \]
\[ \xrightarrow{d} E[(\hat{f}'_1\gamma_1)^2] \left( 1 + \frac{\hat{u}^2}{w'w} \right) \left[ (\hat{B}'\hat{B})^{-1} + \frac{(\hat{B}'\hat{B})^{-1}\hat{B}'zz'(\hat{B}'\hat{B})^{-1}}{w'w} \right] + C \]
\[ + u' \left( I_{N-K} - \frac{ww'}{w'w} \right) u \frac{(\hat{B}'\hat{B})^{-1}\hat{B}'zz'(\hat{B}'\hat{B})^{-1}}{(w'w)^2}. \]  
(A.40)
Let $b_i$ be the $i$-th diagonal element of $(\hat{B}'\hat{B})^{-1}$. Then, we can readily show that
\[ \dot{z}_i = -\frac{\ell_i'(\hat{B}'\hat{B})^{-1}\hat{B}'z}{\sqrt{\hat{b}_i}} \sim N(0,1), \]
\[ \nu = \frac{u'[I_{N-K} - w(u'w)^{-1}u']u}{E[(\hat{f}_i\gamma_i^*)^2]} \sim \chi^2_{N-K-1}, \]
and $\nu$ is independent of $\tilde{u}, z$ and $w$. Using $\dot{z}_i$ and $\nu$, we can express the limiting distribution of $s_m^2(\hat{\gamma}_{1i})$ as
\[ Ts_m^2(\hat{\gamma}_{1i}) = T\ell_i'\hat{V}_m(\hat{\gamma}_1)\ell_i \xrightarrow{d} E[(\hat{f}_i\gamma_i^*)^2]b_i \left(1 + \frac{\tilde{u}^2}{w'w}\right) \left(1 + \frac{\tilde{z}_i^2}{w'w}\right) + \frac{\tilde{z}_i^2\nu}{(w'w)^2} + c_i, \]
where $c_i$ is the $i$-th diagonal element of $C$. In addition, by letting
\[ \tilde{r}_i = (E[(\hat{f}_i\gamma_i^*)^2]b_i + c_i)^{-\frac{1}{2}} \ell_i'(\hat{B}'\hat{B})^{-\frac{1}{2}}r \sim N(0,1), \]
we can write the $i$-th element in (A.23) as
\[ \sqrt{T}(\hat{\gamma}_{1i} - \gamma_{1i}^*) \xrightarrow{d} (E[(\hat{f}_i\gamma_i^*)^2]b_i)^{\frac{1}{2}} \frac{\tilde{u}\dot{z}_i}{(w'w)^{\frac{1}{2}}} + (E[(\hat{f}_i\gamma_i^*)^2]b_i + c_i)^{\frac{1}{2}}\tilde{r}_i. \]

Finally, by letting\(^\text{15}\)
\[ \lambda_i = 1 + \frac{c_i}{E[(\hat{f}_i\gamma_i^*)^2]b_i} > 0, \]
we can write the limiting distribution of $t_m(\hat{\gamma}_{1i})$ as
\[ t_m(\hat{\gamma}_{1i}) = \frac{\hat{\gamma}_{1i} - \gamma_{1i}^*}{s_m(\hat{\gamma}_{1i})} \xrightarrow{d} \frac{\tilde{u}\dot{z}_i + \sqrt{\lambda_i}\sqrt{w'w}\tilde{r}_i}{[\lambda_i(w'w) + \tilde{z}_i^2 + \tilde{u}^2\left(1 + \frac{\tilde{z}_i^2}{w'w}\right) + \frac{\tilde{u}^2\nu}{(w'w)^2}]^{\frac{1}{2}}}. \]
The estimated covariance matrix of $\hat{\gamma}_1$ that assumes a correctly specified model is obtained by dropping the second term in (A.40). Then, it can be shown that
\[ Ts_c^2(\hat{\gamma}_{1i}) \xrightarrow{d} E[(\hat{f}_i\gamma_i^*)^2]b_i \left(1 + \frac{\tilde{u}^2}{w'w}\right) \left(1 + \frac{\tilde{z}_i^2}{w'w}\right) + c_i \]
and hence
\[ t_c(\hat{\gamma}_{1i}) = \frac{\hat{\gamma}_{1i} - \gamma_{1i}^*}{s_c(\hat{\gamma}_{1i})} \xrightarrow{d} \frac{\tilde{u}\dot{z}_i + \sqrt{\lambda_i}\sqrt{w'w}\tilde{r}_i}{[\lambda_i(w'w) + \tilde{z}_i^2 + \tilde{u}^2\left(1 + \frac{\tilde{z}_i^2}{w'w}\right) + \frac{\tilde{u}^2\nu}{(w'w)^2}]^{\frac{1}{2}}}. \]

We now turn our attention to the limiting distributions of $t_c(\hat{\gamma}_2)$ and $t_m(\hat{\gamma}_2)$. From part (a) of Theorem 1, we have
\[ \hat{\gamma}_2 \xrightarrow{d} \frac{w'u}{w'w} = \left(\frac{w'V_{uu}w}{w'w}\right)^{\frac{1}{2}}\tilde{u}, \]
\[ \text{From (A.44), we can see that } E[(\hat{f}_i\gamma_i^*)^2]b_i + c_i \text{ is the variance of } \ell_i'(\hat{B}'\hat{B})^{-\frac{1}{2}}r. \text{ Therefore, we have } \lambda_i > 0. \]

33
where \( \tilde{u} = w'u/(w'V_u w)^{1/2} \sim N(0, 1) \), and it is independent of \( w \). Using (A.37), we obtain

\[
\begin{align*}
\hat{s}^2_m(\hat{\gamma}_2) &= \frac{1}{T^2} \sum_{t=1}^T \hat{h}^2_{2t} \\
& \xrightarrow{d} \frac{1}{(w'u)^2} \left[ w'V_u w + \frac{(w'u)^2}{w'w} \right] + \frac{u'[I_{N-K} - w(w'w)^{-1}w']u}{(w'w)^2} \\
&= \frac{w'V_u w + u'u}{(w'w)^2}. \quad (A.51)
\end{align*}
\]

Therefore, the \( t \)-statistic of \( \hat{\gamma}_2 \) under the misspecification-robust standard error is given by

\[
t_m(\hat{\gamma}_2) = \frac{\hat{\gamma}_2}{s_m(\hat{\gamma}_2)} \xrightarrow{d} \frac{\tilde{u}}{\left( 1 + \frac{w'u}{w'w} \right)^{1/2}}. \quad (A.52)
\]

For \( s^2_c(\hat{\gamma}_2) \) which assumes a correctly specified model, we drop the second term in \( \hat{h}_{2t} \), and we obtain

\[
\begin{align*}
\hat{s}^2_c(\hat{\gamma}_2) &\xrightarrow{d} \frac{1}{(w'u)^2} \left[ w'V_u w + \frac{(w'u)^2}{w'w} \right] = \frac{w'V_u w}{(w'w)^2} \left( 1 + \frac{\tilde{u}^2}{w'w} \right). \quad (A.53)
\end{align*}
\]

It follows that

\[
t_c(\hat{\gamma}_2) = \frac{\hat{\gamma}_2}{s_c(\hat{\gamma}_2)} \xrightarrow{d} \frac{\tilde{u}}{\left( 1 + \frac{\tilde{u}^2}{w'w} \right)^{1/2}}. \quad (A.54)
\]

Under the conditional homoskedasticity assumption, \( V_u = E[(\hat{\gamma}_1^*)^2]I_{N-K} \), so we can write

\[
t_m(\hat{\gamma}_2) \xrightarrow{d} \frac{\tilde{u}}{\left( 1 + \frac{\tilde{u}^2+v}{w'w} \right)^{1/2}}, \quad (A.55)
\]

where \( v \) is defined in (A.42). This completes the proof of part (a) of Theorem 2.

**part (b):** We first derive the limiting distribution of \( \hat{h}_t \) in (A.25). When a model is misspecified, we can see from part (b) of Theorem 1 that \( \hat{\gamma}_2 = O_p(T^{1/2}) \) and \( \hat{\gamma}_1 = O_p(1) \), so \( \hat{\gamma}_2 \) is the dominant term. Therefore, using (28), we have

\[
\hat{e}_t = x_t(\hat{P}'\hat{\gamma}_1 + g_t\hat{\gamma}_2) - q = x_tg_t\hat{\gamma}_2 + O_p(1) = \frac{\sqrt{T}\delta s}{w'w}x_tg_t + O_p(1). \quad (A.56)
\]

In addition, using (A.31), (A.32) and (A.18), we have

\[
-\hat{U}^{-1/2}\hat{e} = \hat{M}\hat{U}^{-1/2}q - \hat{M}\hat{U}^{-1/2}\hat{\gamma}_2 \Rightarrow Mq - \frac{Mz'M\hat{q}}{z'Mz} = P[I_{N-K} - w(w'w)^{-1}w']P'q. \quad (A.57)
\]

It follows that under a misspecified model,

\[
\hat{u}_t = \hat{e}'\hat{U}^{-1}x_t \xrightarrow{d} -\bar{q}'P[I_{N-K} - w(w'w)^{-1}w']P'U^{-1/2}x_t. \quad (A.58)
\]
Then, using (A.28) and (A.30), we can express the limiting distribution of \( \hat{h}_t = [\hat{h}_{1t}', \hat{h}_{2t}'] \) as

\[
\frac{\hat{h}_{1t}}{\sqrt{T}} \overset{d}{\rightarrow} \frac{q' P w}{w' w} (\tilde{B}' \tilde{B})^{-1} \tilde{B}' \left( I_N - \frac{z w'}{w' w} P' \right) U^{-\frac{1}{2}} x_t g_t \\
+ (\tilde{B}' \tilde{B})^{-1} (\tilde{B}' z) q' P[I_{N-K} - w(w' w)^{-1} w'] P' U^{-\frac{1}{2}} x_t g_t,
\]

(A.59)

\[
\frac{\hat{h}_{2t}}{T} \overset{d}{\rightarrow} \frac{q' P w}{(w' w)^2} w' P' U^{-\frac{1}{2}} x_t g_t - \frac{1}{w' w} q' P[I_{N-K} - w(w' w)^{-1} w'] P' U^{-\frac{1}{2}} x_t g_t.
\]

(A.60)

Using the fact that \( P' \tilde{B} = 0_{(N-K) \times K} \) and \( [I_{N-K} - w(w' w)^{-1} w'] w = 0_{N-K} \), we have

\[
\tilde{B}' \left( I_N - \frac{z w'}{w' w} P' \right) P[I_{N-K} - w(w' w)^{-1} w'] P' \tilde{q} = 0_K,
\]

(A.61)

and we can show that the two terms in the limiting distribution of \( \hat{h}_{1t}/\sqrt{T} \) are asymptotically uncorrelated. It follows that

\[
\hat{V}_m(\gamma_1) = \frac{1}{T^2} \sum_{t=1}^T \hat{h}_{1t} \hat{h}_{1t}'
= \frac{(q' P w)^2}{(w' w)^2} \left[ (\tilde{B}' \tilde{B})^{-1} + \frac{(\tilde{B}' \tilde{B})^{-1} \tilde{B}' zz' \tilde{B}(\tilde{B}' \tilde{B})^{-1}}{w' w} \right]
+ \frac{1}{(w' w)^2} \left[ q' P P' \tilde{q} - \frac{(q' P w)^2}{w' w} \right] (\tilde{B}' \tilde{B})^{-1} \tilde{B}' zz' \tilde{B}(\tilde{B}' \tilde{B})^{-1}
= \frac{\delta^2}{(w' w)^2} \left[ s^2 (\tilde{B}' \tilde{B})^{-1} + (\tilde{B}' \tilde{B})^{-1} \tilde{B}' zz' \tilde{B}(\tilde{B}' \tilde{B})^{-1} \right].
\]

(A.62)

Using \( \tilde{z}_i \) as defined in (A.41), we can express the limiting distribution of \( s_m^2(\gamma_1) \) as

\[
s_m^2(\gamma_1) = \frac{\delta^2 b_i}{(w' w)^2} (s^2 + \tilde{z}_i^2).
\]

(A.63)

In addition, we can also use \( \tilde{z}_i \) to express the \( i \)-th element in (27) as

\[
\tilde{\gamma}_i = \gamma_i + \frac{\delta s \sqrt{h_i} \tilde{z}_i}{w' w}.
\]

(A.64)

It follows that when the model is misspecified, \( t_m(\gamma_1) \) has the following limiting distribution:

\[
t_m(\gamma_1) = \frac{\tilde{\gamma}_i - \gamma_i}{s_m(\gamma_1)} \overset{d}{\rightarrow} \frac{s \tilde{z}_i}{\sqrt{s^2 + \tilde{z}_i^2}}.
\]

(A.65)

To show that \( t_m(\gamma_1) \overset{d}{\rightarrow} N(0, 1/4) \), consider the polar transformation \( s = \omega \cos(\theta) \) and \( \tilde{z}_i = \omega \sin(\theta) \), where \( \omega = \sqrt{s^2 + \tilde{z}_i^2} \). The joint density of \( (\omega, \theta) \) is given by

\[
f(\omega, \theta) = \frac{\omega e^{-\frac{\omega^2}{2}}}{2\pi} I_{(\omega > 0)} I_{(0 < \theta < 2\pi)}.
\]

(A.66)
Therefore, $\omega$ and $\theta$ are independent. Using the polar transformation, we obtain

$$\frac{s\tilde{z}_i}{\sqrt{s^2 + \tilde{z}_i^2}} = \omega \cos(\theta) \sin(\theta) = \frac{\omega \sin(2\theta)}{2}.$$  \hspace{1cm} (A.67)

Since $\theta$ is uniformly distributed over $(0, 2\pi)$, $\sin(\theta)$ and $\sin(2\theta)$ have the same distribution. It follows that $\omega \sin(2\theta) \sim N(0, 1)$. Therefore,

$$t_m(\hat{\gamma}_{1i}) \overset{d}{\to} N\left(0, \frac{1}{4}\right).$$  \hspace{1cm} (A.68)

The estimated covariance matrix of $\hat{\gamma}_i$ that assumes a correctly specified model is obtained by dropping the second term in the line before (A.62). We can then show that

$$s^2_c(\hat{\gamma}_{1i}) \overset{d}{\to} \frac{\delta^2 s^2 b_i}{(w'w)^2} \left(1 + \frac{\tilde{z}_i^2}{w'w}\right).$$  \hspace{1cm} (A.69)

Using (A.64), we can then obtain the limiting distribution of $t_c(\hat{\gamma}_{1i})$ as

$$t_c(\hat{\gamma}_{1i}) = \frac{\hat{\gamma}_{1i} - \gamma^*_1}{s_c(\hat{\gamma}_{1i})} \overset{d}{\to} \tilde{z}_i \left(1 + \frac{\tilde{z}_i^2}{w'w}\right)^{-\frac{1}{2}}.$$  \hspace{1cm} (A.70)

Turning our attention to the limiting distributions of $t_c(\hat{\gamma}_{2i})$ and $t_m(\hat{\gamma}_{2i})$, we use (A.60) and the fact that $\delta^2 = \tilde{q}'PP\tilde{q}$ to obtain

$$\frac{s^2_m(\hat{\gamma}_2)}{T} = \frac{1}{T} \sum_{t=1}^{T} \hat{h}_{2t}^2$$

$$\overset{d}{\to} \frac{(\tilde{q}'P\tilde{w})^2}{(w'w)^2} w'w + \frac{1}{(w'w)^2} \tilde{q}'P \left(I_{N-K} - \frac{ww'}{w'w}\right) P'\tilde{q}$$

$$= \frac{\delta^2}{(w'w)^2}.$$  \hspace{1cm} (A.71)

Therefore, using (28), the $t$-statistic of $\hat{\gamma}_2$ under the misspecification-robust standard error is given by

$$t_m(\hat{\gamma}_2) = \frac{\hat{\gamma}_2}{s_m(\hat{\gamma}_2)} \overset{d}{\to} s \sim N(0, 1).$$  \hspace{1cm} (A.72)

For $s^2_c(\hat{\gamma}_2)$ which assumes a correctly specified model, we drop the second term of $\hat{h}_{2t}$ in (A.60), and we obtain

$$\frac{s^2_c(\hat{\gamma}_2)}{T} \overset{d}{\to} \frac{(\tilde{q}'P\tilde{w})^2}{(w'w)^3} = \frac{\delta^2 s^2}{(w'w)^3}.$$  \hspace{1cm} (A.73)
It follows that
\[ t_c(\hat{\gamma}_2) = \frac{\hat{\gamma}_2}{s_{c}(\hat{\gamma}_2)} \xrightarrow{d} \text{sgn}(s)\sqrt{w'w}. \] (A.74)

Note that since \( s \sim N(0,1) \), \( \text{sgn}(s) \) has probabilities of \( 1/2 \) of taking the values of \( -1 \) or \( 1 \), and it is independent of \( s^2 \). As a result, \( \text{sgn}(s) \) is also independent of \( w'w \sim \chi^2_{N-K} \). This completes the proof of part (b) of Theorem 2.

Proof of Corollary 1.

We only provide the proof of part (a) since the proof of part (b) is similar for \( t^2_m(\hat{\gamma}_1) \) and obvious for \( t^2_n(\hat{\gamma}_1) \). First, comparing the limiting distribution of \( t^2_c(\hat{\gamma}_1) \) with the limiting distribution of \( t^2_n(\hat{\gamma}_1) \) in part (a) of Theorem 2, we see that there is an extra positive term \( \tilde{z}_i^2 v/(w'w) \) in the denominator. Therefore, the limiting distribution of \( t^2_n(\hat{\gamma}_1) \) is stochastically dominated by the limiting distribution of \( t^2_c(\hat{\gamma}_1) \). It remains to be shown that the latter is stochastically dominated by \( \chi^2_1 \). From part (a) of Theorem 2, we have
\[ t^2_c(\hat{\gamma}_1) \xrightarrow{d} \frac{(\tilde{z}_i^2 + \sqrt{\lambda_i \sqrt{w'w}})^2}{\lambda_i (w'w) + \tilde{z}_i^2 + \tilde{u}^2 (1 + \frac{\tilde{z}_i^2}{w'w})}. \] (A.76)

Let \( \tilde{t} = \tilde{z}_i/\sqrt{w'w} \). It is easy to see that the limit of \( t^2_c(\hat{\gamma}_1) \) is stochastically dominated by \( (\tilde{t}\tilde{u} + \sqrt{\lambda_i \tilde{r}_i})^2 / (\lambda_i + \tilde{u}^2) \sim \chi^2_1 \).

Next, since \( 1 + \tilde{u}^2/(w'w) > 1 \) and \( 1 + (\tilde{u}^2 + v)/(w'w) > 1 \) almost surely, both the limiting distributions of \( t^2_c(\hat{\gamma}_2) \) and \( t^2_m(\hat{\gamma}_2) \) are stochastically dominated by \( \tilde{u}^2 \sim \chi^2_1 \). This completes the proof of Corollary 1.

Proof of Theorem 3.

**part (a):** Using (A.33) in the proof of Theorem 2, we can easily obtain
\[ T\hat{\delta}^2 = T\hat{e}'\hat{U}^{-1}\hat{e} \xrightarrow{d} \text{u}'[I_{N-K} - w(w'w)^{-1}w']\text{u} = \text{u}'P_wP'_w\text{u}, \] (A.77)

\footnote{It is straightforward to show that the limiting probability density function of \( t_c(\hat{\gamma}_2) \) is
\[ f(t) = \frac{|t|^{N-K-1}e^{-\frac{t^2}{2}}}{2^{\frac{N-K}{2}}\Gamma(\frac{N-K}{2})}. \] (A.75)}
where $P_w$ is an $(N - K) \times (N - K - 1)$ orthonormal matrix such that $P_wP_w' = I_{N-K} - w(w'w)^{-1}w'$. Let $\tilde{v} = (P_w'V_uP_w)^{-\frac12}P_w'u \sim N(0_{N-K-1}, I_{N-K-1})$, which is independent of $w$. Then, we have

$$T\hat{\delta}^2 \overset{d}{\to} \tilde{v}'(P_w'V_uP_w)\tilde{v}. \quad (A.78)$$

For testing $H_0 : \delta = 0$, $T\hat{\delta}^2$ is compared with $\sum_{i=1}^{N-K-1} \hat{\xi}_i X_i$, where the $X_i$’s are independent chi-squared random variables with one degree of freedom and the $\hat{\xi}_i$’s are the $N - K - 1$ nonzero eigenvalues of

$$\hat{S}_1^{-\frac12}\hat{U}^{-\frac12}S_1^{-\frac12} = \hat{S}_1^{-\frac12}\hat{U}^{-\frac12}D(\hat{D}'\hat{U}^{-1}\hat{D})^{-1}\hat{D}'\hat{U}^{-1}\hat{S}_1^{-\frac12}. \quad (A.79)$$

Using (A.32), we can write the above matrix as

$$\hat{S}_1^{-\frac12}\hat{U}^{-\frac12}[I_N - \hat{U}^{-\frac12}\hat{D}(\hat{D}'\hat{U}^{-1}\hat{D})^{-1}\hat{D}'\hat{U}^{-\frac12}]\hat{U}^{-\frac12}\hat{S}_1^{-\frac12}$$

$$= \hat{S}_1^{-\frac12}\hat{U}^{-\frac12}\hat{M}\hat{U}^{-\frac12}\hat{S}_1^{-\frac12} - \hat{S}_1^{-\frac12}\hat{U}^{-\frac12}\hat{M}\hat{U}^{-\frac12}\hat{d}(\hat{d}'\hat{M}\hat{U}^{-\frac12}\hat{d})^{-1}\hat{d}'\hat{U}^{-\frac12}\hat{M}\hat{U}^{-\frac12}\hat{S}_1^{-\frac12}. \quad (A.80)$$

Let $\hat{P}$ be an $N \times (N-K)$ orthonormal matrix such that $\hat{P}\hat{P}' = \hat{M}$ and $\hat{P}_w$ be an $(N-K) \times (N-K-1)$ orthonormal matrix such that $\hat{P}_w\hat{P}_w' = I_{N-K} - \hat{P}'U^{-\frac12}\hat{d}(\hat{d}'\hat{M}\hat{U}^{-\frac12}\hat{d})^{-1}\hat{d}'\hat{U}^{-\frac12}\hat{P}$. We can easily show that $\hat{\xi}_i$’s are the nonzero eigenvalues of

$$\hat{S}_1^{-\frac12}\hat{U}^{-\frac12}\hat{P}\hat{P}_w'\hat{P}'\hat{U}^{-\frac12}\hat{S}_1^{-\frac12}, \quad (A.81)$$

or equivalently the eigenvalues of

$$\hat{P}_w'\hat{P}'\hat{U}^{-\frac12}\hat{S}\hat{U}^{-\frac12}\hat{P}\hat{P}_w. \quad (A.82)$$

Using (A.35), we can show that

$$\hat{P}'\hat{U}^{-\frac12}\hat{e}_t \overset{d}{\to} P'U^{-\frac12}e_t(\gamma'\gamma) + \frac{w'u}{w'w}P'U^{-\frac12}x_t\gamma. \quad (A.83)$$

It follows that

$$\hat{P}'\hat{U}^{-\frac12}\hat{S}\hat{U}^{-\frac12}\hat{P} \overset{d}{\to} P'U^{-\frac12}\hat{S}U^{-\frac12}P + \frac{(w'u)^2}{(w'w)^2}I_{N-K} = V_u + \frac{(w'V_uw)^2}{(w'w)^2}I_{N-K}, \quad (A.84)$$

where $\tilde{u} = w'u/(w'V_uw)^{1/2} \sim N(0, 1)$ and it is independent of $w$.

Under the conditional homoskedasticity assumption, we have $V_u = E[(\hat{f}'\gamma_1)^2]I_{N-K}$ and hence

$$T\hat{\delta}^2 \overset{d}{\to} E[(\hat{f}'\gamma_1)^2]\tilde{v}'\tilde{v} \sim E[(\hat{f}'\gamma_1)^2]I_{N-K-1}, \quad (A.85)$$

$$\hat{P}_w'\hat{P}'\hat{U}^{-\frac12}\hat{S}\hat{U}^{-\frac12}\hat{P}\hat{P}_w \overset{d}{\to} E[(\hat{f}'\gamma_1)^2] \left(1 + \frac{\tilde{u}^2}{w'w}\right)I_{N-K-1}. \quad (A.86)$$
It follows that
\[ \hat{\xi}_i \xrightarrow{d} E[(\hat{f}'_i \gamma^*)^2] \left(1 + \frac{\hat{u}^2}{w'w}\right) = \frac{E[(\hat{f}'_i \gamma^*)^2]}{Q_1}, \] (A.87)
where \( Q_1 = w'w/(\hat{u}^2 + w'w) \sim \text{Beta} \left( \frac{N-K}{2}, \frac{1}{2} \right) \) and it is independent of \( \hat{v} \). Therefore, the limiting probability of rejection of the HJ-distance test of size \( \alpha \) is
\[ \int_0^1 P \left[ \chi^2_{N-K-1} > \frac{c_\alpha}{q} \right] f_{Q_1}(q) dq, \] (A.88)
where \( c_\alpha \) is the \( 100(1 - \alpha) \) percentile of \( \chi^2_{N-K-1} \). Since \( 0 < Q_1 < 1 \), the limiting probability of rejection is less than \( \alpha \). This completes the proof of part (a) of Theorem 3.

**Part (b):** Using (A.57), the limiting distribution of the squared sample HJ-distance \( \hat{\delta}^2 = \hat{e}'\hat{U}^{-\frac{1}{2}}\hat{e} \) can be obtained as
\[ \hat{\delta}^2 \xrightarrow{d} \hat{q}'P[I_{N-K} - w(w'w)^{-1}w']P\hat{q} \]
\[ = \frac{(q'PP'q)w'[I_{N-K} - P'qPqPP'q)^{-1}q'P]w}{w'w} = \delta^2 Q_2, \] (A.89)
where
\[ Q_2 = \frac{w'[I_{N-K} - P'qPqPP'q)^{-1}q'P]w}{w'w} \sim \text{Beta} \left( \frac{N-K-1}{2}, \frac{1}{2} \right) \] (A.90)
and it is independent of \( w \).

From the proof of part (a), we know that the \( \hat{\xi}_i \)'s are the eigenvalues of
\[ \hat{P}_w'\hat{P}'\hat{U}^{-\frac{1}{2}}\hat{S}\hat{U}^{-\frac{1}{2}}\hat{P}\hat{P}_w. \] (A.91)

From (36) and (28), we have
\[ \frac{\hat{S}}{T} \xrightarrow{d} \frac{\delta^2 s^2}{(w'w)^2} U, \] (A.92)
which implies
\[ \frac{\hat{P}_w'\hat{P}'\hat{U}^{-\frac{1}{2}}\hat{S}\hat{U}^{-\frac{1}{2}}\hat{P}\hat{P}_w}{T} \xrightarrow{d} \frac{\delta^2 s^2}{(w'w)^2} I_{N-K-1} \] (A.93)
and
\[ \frac{\hat{\xi}_i}{T} \xrightarrow{d} \frac{\delta^2 s^2}{(w'w)^2} = \frac{\delta^2(1 - Q_2)}{w'w}. \] (A.94)

When we compare \( T\hat{\delta}^2 \) with the distribution of \( \sum_{i=1}^{N-K-1} \hat{\xi}_i X_i \), we are effectively comparing \( Q_2 \) with \( (1 - Q_2)/(w'w)\chi^2_{N-K-1} \), and we will reject \( H_0 : \delta = 0 \) when
\[ w'w > \frac{c_\alpha Q_2}{1 - Q_2}. \] (A.95)

39
Note that $w'w \sim \chi^2_{N-K}$ and it is independent of $Q_2$, so the limiting probability of rejection for a test with size $\alpha$ is

$$\int_0^1 P \left[ \chi^2_{N-K} > \frac{c\alpha q}{1-q} \right] f_{Q_2}(q) dq. \quad (A.96)$$

This completes the proof of part (b) of Theorem 3.
References


Table 1
Empirical size of the $t$-tests in a model with a useful factor

Panel A: Correctly specified model

<table>
<thead>
<tr>
<th>$t$-test</th>
<th>$T$</th>
<th>$\hat{\gamma}_0$ 10%</th>
<th>$\hat{\gamma}_0$ 5%</th>
<th>$\hat{\gamma}_0$ 1%</th>
<th>$\hat{\gamma}_1$ 10%</th>
<th>$\hat{\gamma}_1$ 5%</th>
<th>$\hat{\gamma}_1$ 1%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_c$</td>
<td>200</td>
<td>0.168 0.138 0.106</td>
<td>0.097 0.049 0.008</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>600</td>
<td>0.134 0.095 0.059</td>
<td>0.100 0.049 0.009</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1,000</td>
<td>0.119 0.078 0.039</td>
<td>0.099 0.049 0.010</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\infty$</td>
<td>0.100 0.050 0.010</td>
<td>0.100 0.050 0.010</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t_m$</td>
<td>200</td>
<td>0.168 0.138 0.106</td>
<td>0.097 0.048 0.008</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>600</td>
<td>0.134 0.095 0.059</td>
<td>0.100 0.049 0.009</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1,000</td>
<td>0.119 0.078 0.039</td>
<td>0.099 0.049 0.010</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\infty$</td>
<td>0.100 0.050 0.010</td>
<td>0.100 0.050 0.010</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel B: Misspecified model

<table>
<thead>
<tr>
<th>$t$-test</th>
<th>$T$</th>
<th>$\hat{\gamma}_0$ 10%</th>
<th>$\hat{\gamma}_0$ 5%</th>
<th>$\hat{\gamma}_0$ 1%</th>
<th>$\hat{\gamma}_1$ 10%</th>
<th>$\hat{\gamma}_1$ 5%</th>
<th>$\hat{\gamma}_1$ 1%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_c$</td>
<td>200</td>
<td>0.167 0.137 0.105</td>
<td>0.099 0.048 0.009</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>600</td>
<td>0.134 0.096 0.059</td>
<td>0.099 0.050 0.010</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>0.122 0.079 0.040</td>
<td>0.100 0.051 0.010</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\infty$</td>
<td>0.100 0.050 0.010</td>
<td>0.100 0.050 0.010</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t_m$</td>
<td>200</td>
<td>0.167 0.137 0.105</td>
<td>0.098 0.048 0.009</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>600</td>
<td>0.134 0.096 0.059</td>
<td>0.099 0.050 0.010</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>0.122 0.078 0.040</td>
<td>0.100 0.051 0.010</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\infty$</td>
<td>0.100 0.050 0.010</td>
<td>0.100 0.050 0.010</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The table presents the empirical size of the $t$-tests of $H_0 : \gamma_i = \gamma_i^*$ ($i = 0, 1$) in a model with a constant and a useful factor. $\gamma_0$ is the coefficient on the constant term and $\gamma_1$ is the coefficient on the useful factor. $t_c$ denotes the $t$-test constructed under the assumption of correct model specification and $t_m$ denotes the misspecification-robust $t$-test. We report results for different levels of significance (10%, 5% and 1%) and for different values of the number of time series observations ($T$) using 100,000 simulations, assuming that the returns are generated from a multivariate normal distribution with means and covariance matrix calibrated to the 25 Fama-French portfolio returns and the risk-free rate for the period 1959:2–2007:7. The various $t$-statistics are compared to the critical values from a standard normal distribution.
Table 2
Empirical size of the $t$-tests in a model with a useless factor

Panel A: Correctly specified model

<table>
<thead>
<tr>
<th>$t$-test</th>
<th>$T$</th>
<th>$\gamma_0$</th>
<th></th>
<th></th>
<th>$\gamma_1$</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>10% 5% 1%</td>
<td>10% 5% 1%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t_c$</td>
<td>200</td>
<td>0.003 0.001 0.000</td>
<td>0.103 0.045 0.005</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>600</td>
<td>0.001 0.000 0.000</td>
<td>0.087 0.036 0.003</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>0.001 0.000 0.000</td>
<td>0.084 0.033 0.003</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\infty$</td>
<td>0.001 0.000 0.000</td>
<td>0.080 0.032 0.003</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t_m$</td>
<td>200</td>
<td>0.001 0.000 0.000</td>
<td>0.026 0.007 0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>600</td>
<td>0.000 0.000 0.000</td>
<td>0.019 0.005 0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>0.000 0.000 0.000</td>
<td>0.017 0.004 0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\infty$</td>
<td>0.000 0.000 0.000</td>
<td>0.016 0.004 0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel B: Misspecified model

<table>
<thead>
<tr>
<th>$t$-test</th>
<th>$T$</th>
<th>$\gamma_0$</th>
<th></th>
<th></th>
<th>$\gamma_1$</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>10% 5% 1%</td>
<td>10% 5% 1%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t_c$</td>
<td>200</td>
<td>0.012 0.003 0.000</td>
<td>0.314 0.212 0.074</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>600</td>
<td>0.017 0.004 0.000</td>
<td>0.484 0.389 0.215</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>0.022 0.005 0.000</td>
<td>0.569 0.485 0.317</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\infty$</td>
<td>0.080 0.032 0.003</td>
<td>1.000 1.000 1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t_m$</td>
<td>200</td>
<td>0.001 0.000 0.000</td>
<td>0.067 0.028 0.003</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>600</td>
<td>0.001 0.000 0.000</td>
<td>0.079 0.035 0.005</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>0.001 0.000 0.000</td>
<td>0.085 0.040 0.006</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\infty$</td>
<td>0.001 0.000 0.000</td>
<td>0.100 0.050 0.010</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The table presents the empirical size of the $t$-tests of $H_0 : \gamma_i = \gamma_i^*$ ($i = 0, 1$) in a model with a constant and a useless factor. $\gamma_0$ is the coefficient on the constant term and $\gamma_1$ is the coefficient on the useless factor. $t_c$ denotes the $t$-test constructed under the assumption of correct model specification and $t_m$ denotes the misspecification-robust $t$-test. We report results for different levels of significance (10%, 5% and 1%) and for different values of the number of time series observations ($T$) using 100,000 simulations, assuming that the returns are generated from a multivariate normal distribution with means and covariance matrix calibrated to the 25 Fama-French portfolio returns and the risk-free rate for the period 1959:2–2007:7. The various $t$-statistics are compared to the critical values from a standard normal distribution. The rejection rates for the limiting case ($T = \infty$) are based on the asymptotic distributions given in Theorem 2.
Table 3
Empirical size of the $t$-tests in a model with a useful and a useless factor

Panel A: Correctly specified model

<table>
<thead>
<tr>
<th>$t$-test</th>
<th>$T$</th>
<th>$\hat{\gamma}_0$</th>
<th></th>
<th></th>
<th></th>
<th>$\hat{\gamma}_1$</th>
<th></th>
<th></th>
<th></th>
<th>$\hat{\gamma}_2$</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>10%</td>
<td>5%</td>
<td>1%</td>
<td></td>
<td>10%</td>
<td>5%</td>
<td>1%</td>
<td></td>
<td>10%</td>
<td>5%</td>
<td>1%</td>
<td></td>
</tr>
<tr>
<td>$t_c$</td>
<td>200</td>
<td>0.033</td>
<td>0.013</td>
<td>0.005</td>
<td></td>
<td>0.090</td>
<td>0.043</td>
<td>0.007</td>
<td></td>
<td>0.101</td>
<td>0.045</td>
<td>0.005</td>
<td></td>
</tr>
<tr>
<td></td>
<td>600</td>
<td>0.040</td>
<td>0.016</td>
<td>0.005</td>
<td></td>
<td>0.092</td>
<td>0.044</td>
<td>0.008</td>
<td></td>
<td>0.086</td>
<td>0.035</td>
<td>0.003</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>0.040</td>
<td>0.017</td>
<td>0.004</td>
<td></td>
<td>0.092</td>
<td>0.045</td>
<td>0.008</td>
<td></td>
<td>0.083</td>
<td>0.034</td>
<td>0.003</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\infty$</td>
<td>0.040</td>
<td>0.014</td>
<td>0.002</td>
<td></td>
<td>0.093</td>
<td>0.045</td>
<td>0.008</td>
<td></td>
<td>0.079</td>
<td>0.031</td>
<td>0.002</td>
<td></td>
</tr>
<tr>
<td>$t_m$</td>
<td>200</td>
<td>0.018</td>
<td>0.008</td>
<td>0.003</td>
<td></td>
<td>0.085</td>
<td>0.040</td>
<td>0.006</td>
<td></td>
<td>0.026</td>
<td>0.007</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>600</td>
<td>0.024</td>
<td>0.011</td>
<td>0.003</td>
<td></td>
<td>0.086</td>
<td>0.040</td>
<td>0.007</td>
<td></td>
<td>0.019</td>
<td>0.004</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>0.026</td>
<td>0.011</td>
<td>0.003</td>
<td></td>
<td>0.086</td>
<td>0.041</td>
<td>0.007</td>
<td></td>
<td>0.018</td>
<td>0.005</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\infty$</td>
<td>0.027</td>
<td>0.010</td>
<td>0.001</td>
<td></td>
<td>0.087</td>
<td>0.041</td>
<td>0.007</td>
<td></td>
<td>0.016</td>
<td>0.004</td>
<td>0.000</td>
<td></td>
</tr>
</tbody>
</table>

Panel B: Misspecified model

<table>
<thead>
<tr>
<th>$t$-test</th>
<th>$T$</th>
<th>$\hat{\gamma}_0$</th>
<th></th>
<th></th>
<th></th>
<th>$\hat{\gamma}_1$</th>
<th></th>
<th></th>
<th></th>
<th>$\hat{\gamma}_2$</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>10%</td>
<td>5%</td>
<td>1%</td>
<td></td>
<td>10%</td>
<td>5%</td>
<td>1%</td>
<td></td>
<td>10%</td>
<td>5%</td>
<td>1%</td>
<td></td>
</tr>
<tr>
<td>$t_c$</td>
<td>200</td>
<td>0.036</td>
<td>0.012</td>
<td>0.003</td>
<td></td>
<td>0.091</td>
<td>0.044</td>
<td>0.007</td>
<td></td>
<td>0.303</td>
<td>0.205</td>
<td>0.069</td>
<td></td>
</tr>
<tr>
<td></td>
<td>600</td>
<td>0.041</td>
<td>0.013</td>
<td>0.002</td>
<td></td>
<td>0.089</td>
<td>0.043</td>
<td>0.007</td>
<td></td>
<td>0.476</td>
<td>0.380</td>
<td>0.205</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>0.045</td>
<td>0.013</td>
<td>0.002</td>
<td></td>
<td>0.090</td>
<td>0.042</td>
<td>0.007</td>
<td></td>
<td>0.565</td>
<td>0.478</td>
<td>0.305</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\infty$</td>
<td>0.079</td>
<td>0.031</td>
<td>0.002</td>
<td></td>
<td>0.079</td>
<td>0.031</td>
<td>0.002</td>
<td></td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>$t_m$</td>
<td>200</td>
<td>0.011</td>
<td>0.004</td>
<td>0.001</td>
<td></td>
<td>0.077</td>
<td>0.036</td>
<td>0.005</td>
<td></td>
<td>0.065</td>
<td>0.027</td>
<td>0.003</td>
<td></td>
</tr>
<tr>
<td></td>
<td>600</td>
<td>0.007</td>
<td>0.003</td>
<td>0.001</td>
<td></td>
<td>0.065</td>
<td>0.028</td>
<td>0.004</td>
<td></td>
<td>0.078</td>
<td>0.034</td>
<td>0.005</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>0.005</td>
<td>0.002</td>
<td>0.000</td>
<td></td>
<td>0.057</td>
<td>0.023</td>
<td>0.003</td>
<td></td>
<td>0.084</td>
<td>0.039</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\infty$</td>
<td>0.001</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
<td>0.001</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
<td>0.100</td>
<td>0.050</td>
<td>0.010</td>
<td></td>
</tr>
</tbody>
</table>

The table presents the empirical size of the $t$-tests of $H_0: \gamma_i = \gamma_i^*$ ($i = 0, 1, 2$) in a model with a constant, a useful and a useless factor. $\gamma_0$ is the coefficient on the constant term, $\gamma_1$ is the coefficient on the useful factor, and $\gamma_2$ is the coefficient on the useless factor. $t_c$ denotes the $t$-test constructed under the assumption of correct model specification and $t_m$ denotes the misspecification-robust $t$-test. We report results for different levels of significance (10%, 5% and 1%) and for different values of the number of time series observations ($T$) using 100,000 simulations, assuming that the returns are generated from a multivariate normal distribution with means and covariance matrix calibrated to the 25 Fama-French portfolio returns and the risk-free rate for the period 1959:2–2007:7. The various $t$-tests are compared to the critical values from a standard normal distribution. The rejection rates for the limiting case ($T = \infty$) are based on the asymptotic distributions given in Theorem 2.
Table 4
Empirical size and power of the HJ-distance test

<table>
<thead>
<tr>
<th></th>
<th>N = 11</th>
<th></th>
<th>N = 26</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Size</td>
<td>Power</td>
<td>Size</td>
<td>Power</td>
</tr>
<tr>
<td></td>
<td>10%</td>
<td>5%</td>
<td>1%</td>
<td>10%</td>
</tr>
<tr>
<td></td>
<td>5%</td>
<td>1%</td>
<td></td>
<td>5%</td>
</tr>
<tr>
<td></td>
<td>1%</td>
<td></td>
<td>1%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>0.129</td>
<td>0.069</td>
<td>0.298</td>
<td>0.195</td>
</tr>
<tr>
<td></td>
<td>0.017</td>
<td>0.298</td>
<td>0.195</td>
<td>0.070</td>
</tr>
<tr>
<td>600</td>
<td>0.110</td>
<td>0.057</td>
<td>0.607</td>
<td>0.480</td>
</tr>
<tr>
<td></td>
<td>0.012</td>
<td>0.607</td>
<td>0.480</td>
<td>0.254</td>
</tr>
<tr>
<td>1000</td>
<td>0.108</td>
<td>0.056</td>
<td>0.829</td>
<td>0.736</td>
</tr>
<tr>
<td></td>
<td>0.012</td>
<td>0.829</td>
<td>0.736</td>
<td>0.509</td>
</tr>
<tr>
<td>∞</td>
<td>0.100</td>
<td>0.050</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>0.010</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Panel B: Model with a useless factor

<table>
<thead>
<tr>
<th></th>
<th>N = 11</th>
<th></th>
<th>N = 26</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Size</td>
<td>Power</td>
<td>Size</td>
<td>Power</td>
</tr>
<tr>
<td></td>
<td>10%</td>
<td>5%</td>
<td>1%</td>
<td>10%</td>
</tr>
<tr>
<td></td>
<td>5%</td>
<td>1%</td>
<td></td>
<td>5%</td>
</tr>
<tr>
<td></td>
<td>1%</td>
<td></td>
<td>1%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>0.091</td>
<td>0.046</td>
<td>0.334</td>
<td>0.228</td>
</tr>
<tr>
<td></td>
<td>0.010</td>
<td>0.334</td>
<td>0.228</td>
<td>0.092</td>
</tr>
<tr>
<td>600</td>
<td>0.076</td>
<td>0.036</td>
<td>0.692</td>
<td>0.597</td>
</tr>
<tr>
<td></td>
<td>0.007</td>
<td>0.692</td>
<td>0.597</td>
<td>0.395</td>
</tr>
<tr>
<td>1000</td>
<td>0.073</td>
<td>0.035</td>
<td>0.830</td>
<td>0.774</td>
</tr>
<tr>
<td></td>
<td>0.006</td>
<td>0.830</td>
<td>0.774</td>
<td>0.639</td>
</tr>
<tr>
<td>∞</td>
<td>0.070</td>
<td>0.033</td>
<td>0.944</td>
<td>0.931</td>
</tr>
<tr>
<td></td>
<td>0.006</td>
<td>0.944</td>
<td>0.931</td>
<td>0.902</td>
</tr>
</tbody>
</table>

Panel C: Model with a useful and a useless factor

<table>
<thead>
<tr>
<th></th>
<th>N = 11</th>
<th></th>
<th>N = 26</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Size</td>
<td>Power</td>
<td>Size</td>
<td>Power</td>
</tr>
<tr>
<td></td>
<td>10%</td>
<td>5%</td>
<td>1%</td>
<td>10%</td>
</tr>
<tr>
<td></td>
<td>5%</td>
<td>1%</td>
<td></td>
<td>5%</td>
</tr>
<tr>
<td></td>
<td>1%</td>
<td></td>
<td>1%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>0.090</td>
<td>0.047</td>
<td>0.200</td>
<td>0.122</td>
</tr>
<tr>
<td></td>
<td>0.010</td>
<td>0.200</td>
<td>0.122</td>
<td>0.039</td>
</tr>
<tr>
<td>600</td>
<td>0.076</td>
<td>0.037</td>
<td>0.423</td>
<td>0.311</td>
</tr>
<tr>
<td></td>
<td>0.007</td>
<td>0.423</td>
<td>0.311</td>
<td>0.142</td>
</tr>
<tr>
<td>1000</td>
<td>0.074</td>
<td>0.035</td>
<td>0.609</td>
<td>0.502</td>
</tr>
<tr>
<td></td>
<td>0.007</td>
<td>0.609</td>
<td>0.502</td>
<td>0.299</td>
</tr>
<tr>
<td>∞</td>
<td>0.069</td>
<td>0.033</td>
<td>0.926</td>
<td>0.910</td>
</tr>
<tr>
<td></td>
<td>0.006</td>
<td>0.926</td>
<td>0.910</td>
<td>0.875</td>
</tr>
</tbody>
</table>

The table presents the empirical size and power of the HJ-distance test of $H_0 : \delta^2 = 0$. We report results for different levels of significance (10%, 5% and 1%) and for different values of the number of time series observations ($T$) using 100,000 simulations, assuming that the returns are generated from a multivariate normal distribution with means and covariance matrix calibrated to the risk-free rate and the 10 size portfolio returns ($N = 11$) or 25 Fama-French ($N = 26$) portfolio returns for the period 1959:2-2007:7. The HJ-distance statistic $T\hat{\delta}^2$ is compared to the critical values from the weighted chi-squared distribution described in Section 2. The rejection rates for the limiting case ($T = \infty$) are based on the asymptotic results given in Theorem 3.
Table 5
Survival rates of risk factors

Panel A: Correctly specified model

<table>
<thead>
<tr>
<th></th>
<th>Useful factor ($\gamma_1^* \neq 0$)</th>
<th>Useful factor ($\gamma_2^* = 0$)</th>
<th>Useless factor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$t_c(\hat{\gamma}_1)$</td>
<td>$t_m(\hat{\gamma}_1)$</td>
<td>$t_c(\hat{\gamma}_2)$</td>
</tr>
<tr>
<td>200</td>
<td>0.455</td>
<td>0.459</td>
<td>0.063</td>
</tr>
<tr>
<td>600</td>
<td>0.900</td>
<td>0.903</td>
<td>0.055</td>
</tr>
<tr>
<td>1000</td>
<td>0.987</td>
<td>0.988</td>
<td>0.050</td>
</tr>
</tbody>
</table>

Panel B: Misspecified model

<table>
<thead>
<tr>
<th></th>
<th>Useful factor ($\gamma_1^* \neq 0$)</th>
<th>Useful factor ($\gamma_2^* = 0$)</th>
<th>Useless factor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$t_c(\hat{\gamma}_1)$</td>
<td>$t_m(\hat{\gamma}_1)$</td>
<td>$t_c(\hat{\gamma}_2)$</td>
</tr>
<tr>
<td>200</td>
<td>0.432</td>
<td>0.447</td>
<td>0.059</td>
</tr>
<tr>
<td>600</td>
<td>0.835</td>
<td>0.868</td>
<td>0.053</td>
</tr>
<tr>
<td>1000</td>
<td>0.932</td>
<td>0.954</td>
<td>0.049</td>
</tr>
</tbody>
</table>

The table presents the survival rates of the useful and useless factors in a model with a constant, a useful factor (with $\gamma_1^* \neq 0$), a useful factor that does not contribute to pricing (with $\gamma_2^* = 0$) and a useless factor (with $\gamma_3^*$ unidentified). The sequential procedure is implemented by using the misspecification-robust $t$-tests ($t_m(\hat{\gamma}_i)$ column) as well as the $t$-tests under correctly specified models ($t_c(\hat{\gamma}_i)$ column). The nominal level of the sequential testing procedure is set equal to 5%. Panels A and B are for correctly specified and misspecified models, respectively. We report results for different values of the number of time series observations ($T$) using 100,000 simulations, assuming that the returns are generated from a multivariate normal distribution with means and covariance matrix calibrated to the 25 Fama-French portfolio returns and the risk-free rate for the period 1959:2–2007:7.
The table presents the probability that both factors survive, only one factor survives, and no factor survives in a model in which a linear combination of two useful factors is useless. The sequential procedure is implemented by using the misspecification-robust $t$-test ($t_m$ column) as well as the $t$-test under correctly specified models ($t_c$ column). The nominal level of the sequential testing procedure is set equal to 5%. Panels A and B are for correctly specified and misspecified models, respectively. We report results for different values of the number of time series observations ($T$) using 100,000 simulations, assuming that the returns are generated from a multivariate normal distribution with means and covariance matrix calibrated to the 25 Fama-French portfolio returns and the risk-free rate for the period 1959:2–2007:7.

### Panel A: Correctly specified model

<table>
<thead>
<tr>
<th>$T$</th>
<th>$t_c$</th>
<th>$t_m$</th>
<th>$t_c$</th>
<th>$t_m$</th>
<th>$t_c$</th>
<th>$t_m$</th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td>0.035</td>
<td>0.005</td>
<td>0.364</td>
<td>0.372</td>
<td>0.601</td>
<td>0.622</td>
</tr>
<tr>
<td>600</td>
<td>0.028</td>
<td>0.004</td>
<td>0.792</td>
<td>0.811</td>
<td>0.180</td>
<td>0.186</td>
</tr>
<tr>
<td>1000</td>
<td>0.026</td>
<td>0.003</td>
<td>0.933</td>
<td>0.954</td>
<td>0.041</td>
<td>0.042</td>
</tr>
</tbody>
</table>

### Panel B: Misspecified model

<table>
<thead>
<tr>
<th>$T$</th>
<th>$t_c$</th>
<th>$t_m$</th>
<th>$t_c$</th>
<th>$t_m$</th>
<th>$t_c$</th>
<th>$t_m$</th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td>0.186</td>
<td>0.024</td>
<td>0.314</td>
<td>0.371</td>
<td>0.500</td>
<td>0.605</td>
</tr>
<tr>
<td>600</td>
<td>0.358</td>
<td>0.032</td>
<td>0.524</td>
<td>0.789</td>
<td>0.118</td>
<td>0.179</td>
</tr>
<tr>
<td>1000</td>
<td>0.456</td>
<td>0.037</td>
<td>0.521</td>
<td>0.922</td>
<td>0.023</td>
<td>0.041</td>
</tr>
</tbody>
</table>

Table 6
Survival rates when a linear combination of the factors is useless
Table 7
Monthly analysis of some popular linear asset pricing models

Panel A: Rank test for individual factors

<table>
<thead>
<tr>
<th>Test</th>
<th>vw</th>
<th>smb</th>
<th>hml</th>
<th>c</th>
<th>cnd</th>
<th>cd</th>
<th>cay</th>
<th>c-cay</th>
<th>prem</th>
<th>lab</th>
<th>term</th>
<th>def</th>
<th>div</th>
<th>rf</th>
</tr>
</thead>
<tbody>
<tr>
<td>W*</td>
<td>168.4</td>
<td>167.3</td>
<td>156.3</td>
<td>39.4</td>
<td>33.5</td>
<td>37.4</td>
<td>37.3</td>
<td>22.6</td>
<td>32.4</td>
<td>33.6</td>
<td>30.9</td>
<td>38.5</td>
<td>144.8</td>
<td>21.7</td>
</tr>
<tr>
<td>p-val</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.033</td>
<td>0.118</td>
<td>0.053</td>
<td>0.054</td>
<td>0.600</td>
<td>0.148</td>
<td>0.116</td>
<td>0.193</td>
<td>0.041</td>
<td>0.000</td>
<td>0.651</td>
</tr>
</tbody>
</table>

Panel B: HJ-distance, Lagrange multiplier, and rank tests

<table>
<thead>
<tr>
<th>Model</th>
<th>( \hat{\delta} )</th>
<th>p-val</th>
<th>LM p-val</th>
<th>W* p-val</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAPM</td>
<td>0.430</td>
<td>0.000</td>
<td>92.927</td>
<td>0.000</td>
</tr>
<tr>
<td>C-LAB</td>
<td>0.406</td>
<td>0.000</td>
<td>68.034</td>
<td>0.000</td>
</tr>
<tr>
<td>FF3</td>
<td>0.367</td>
<td>0.000</td>
<td>66.358</td>
<td>0.000</td>
</tr>
<tr>
<td>ICAPM</td>
<td>0.339</td>
<td>0.252</td>
<td>23.962</td>
<td>0.244</td>
</tr>
<tr>
<td>CCAPM</td>
<td>0.429</td>
<td>0.000</td>
<td>80.135</td>
<td>0.000</td>
</tr>
<tr>
<td>CC-CAY</td>
<td>0.395</td>
<td>0.001</td>
<td>50.374</td>
<td>0.001</td>
</tr>
<tr>
<td>D-CCAPM</td>
<td>0.408</td>
<td>0.000</td>
<td>70.547</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Panel C: Model selection procedure using standard errors under correct model specification

<table>
<thead>
<tr>
<th>Model</th>
<th>vw</th>
<th>smb</th>
<th>hml</th>
<th>c</th>
<th>cnd</th>
<th>cd</th>
<th>cay</th>
<th>c-cay</th>
<th>prem</th>
<th>lab</th>
<th>term</th>
<th>def</th>
<th>div</th>
<th>rf</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAPM</td>
<td>-2.60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C-LAB</td>
<td>-1.18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FF3</td>
<td>-3.67</td>
<td>-1.77</td>
<td>-5.31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICAPM</td>
<td>-4.19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCAPM</td>
<td>1.56</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC-CAY</td>
<td>-2.22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D-CCAPM</td>
<td>-0.89</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

50
Table 7 (continued)
Monthly analysis of some popular linear asset pricing models

Panel D: Model selection procedure using model misspecification-robust standard errors

<table>
<thead>
<tr>
<th>Model</th>
<th>vw</th>
<th>smb</th>
<th>hml</th>
<th>c</th>
<th>c_ne</th>
<th>c_d</th>
<th>c_cay</th>
<th>c_cay</th>
<th>prem</th>
<th>lab</th>
<th>term</th>
<th>def</th>
<th>div</th>
<th>rf</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAPM</td>
<td>-2.60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C-LAB</td>
<td>-1.03</td>
<td>-0.60</td>
<td>1.17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-1.45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FF3</td>
<td>-3.66</td>
<td>-1.77</td>
<td>-5.33</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-4.20</td>
<td></td>
<td>-4.89</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICAPM</td>
<td>1.16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-1.76</td>
<td>1.53</td>
<td>1.27</td>
<td>0.49</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-2.34</td>
<td>1.46</td>
<td>1.28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-2.33</td>
<td>1.26</td>
<td>1.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-2.67</td>
<td>1.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-2.47</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCAPM</td>
<td>-1.58</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC-CAY</td>
<td>-1.66</td>
<td>-1.84</td>
<td>0.45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-1.64</td>
<td>-1.81</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-1.87</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-1.30</td>
<td>-0.79</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-1.98</td>
<td>-0.67</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-2.12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The table presents the estimation results of seven asset pricing models. The models include the CAPM, the conditional CAPM (C-LAB) of Jagannathan and Wang (1996), the Fama and French (1993) three-factor model (FF3), the intertemporal CAPM (ICAPM) specification of Petkova (2006), the consumption CAPM (CCAPM), the conditional consumption CAPM (CC-CAY) of Lettau and Ludvigson (2001), and the durable consumption CAPM (D-CCAPM) of Yogo (2006). The models are estimated using monthly gross returns on the 25 Fama-French size and book-to-market ranked portfolios and the one-month T-bill. The data are from 1959:2 until 2007:7. Panel A reports the rank restriction test ($W^*$) and its p-value ($p$-val) of the null that $E[x_t(1, f_{it})]$ has a column rank of one. In Panel B, we report the sample HJ-distance ($\hat{\delta}$), the Lagrange multiplier ($LM$) test, and the rank restriction test ($W^*$) with the corresponding p-values ($p$-val) for each model. The $t$-tests of the model selection procedures based on the standard errors under correct model specification and model misspecification are in Panels C and D, respectively. Each $t$-test is for the test of the null hypothesis that the coefficient associated with a given risk factor is equal to zero. We use boldface to highlight those cases in which the factors survive the model selection procedure at the 5% significance level.
Table 8
Quarterly analysis of some popular linear asset pricing models

Panel A: Rank test for individual factors

<table>
<thead>
<tr>
<th>Test</th>
<th>vw</th>
<th>vw\cdot ml</th>
<th>smb</th>
<th>hml</th>
<th>c</th>
<th>c_{nd}</th>
<th>c_{d}</th>
<th>c_{ay}</th>
<th>c_{my}</th>
<th>i_{h}</th>
<th>i_{c}</th>
<th>i_{nc}</th>
</tr>
</thead>
<tbody>
<tr>
<td>W*</td>
<td>74.3</td>
<td>74.4</td>
<td>93.2</td>
<td>68.6</td>
<td>32.6</td>
<td>29.4</td>
<td>40.4</td>
<td>35.3</td>
<td>21.2</td>
<td>32.4</td>
<td>20.2</td>
<td>32.8</td>
</tr>
<tr>
<td>p-val</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.140</td>
<td>0.248</td>
<td>0.026</td>
<td>0.083</td>
<td>0.680</td>
<td>0.147</td>
<td>0.734</td>
<td>0.137</td>
<td>0.416</td>
</tr>
</tbody>
</table>

Panel B: HJ-distance, Lagrange multiplier, and rank tests

<table>
<thead>
<tr>
<th>Model</th>
<th>\hat{\delta}</th>
<th>p-val</th>
<th>LM p-val</th>
<th>W* p-val</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAPM</td>
<td>0.629</td>
<td>0.000</td>
<td>63.904</td>
<td>0.000</td>
</tr>
<tr>
<td>C-ML</td>
<td>0.608</td>
<td>0.000</td>
<td>63.995</td>
<td>0.000</td>
</tr>
<tr>
<td>FF3</td>
<td>0.549</td>
<td>0.000</td>
<td>49.269</td>
<td>0.001</td>
</tr>
<tr>
<td>CCAPM</td>
<td>0.658</td>
<td>0.000</td>
<td>66.084</td>
<td>0.000</td>
</tr>
<tr>
<td>CC-CAY</td>
<td>0.587</td>
<td>0.003</td>
<td>37.559</td>
<td>0.021</td>
</tr>
<tr>
<td>D-CCAPM</td>
<td>0.612</td>
<td>0.000</td>
<td>59.720</td>
<td>0.000</td>
</tr>
<tr>
<td>CC-MY</td>
<td>0.625</td>
<td>0.000</td>
<td>49.750</td>
<td>0.001</td>
</tr>
<tr>
<td>SIM</td>
<td>0.540</td>
<td>0.129</td>
<td>35.155</td>
<td>0.037</td>
</tr>
</tbody>
</table>

Panel C: Model selection procedure using standard errors under correct model specification

<table>
<thead>
<tr>
<th>Model</th>
<th>vw</th>
<th>vw\cdot ml</th>
<th>smb</th>
<th>hml</th>
<th>c</th>
<th>c_{nd}</th>
<th>c_{d}</th>
<th>c_{ay}</th>
<th>c_{my}</th>
<th>i_{h}</th>
<th>i_{c}</th>
<th>i_{nc}</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAPM</td>
<td>-2.94</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C-ML</td>
<td>2.07</td>
<td>-2.10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FF3</td>
<td>-3.69</td>
<td>-0.28</td>
<td>-4.33</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCAPM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC-CAY</td>
<td></td>
<td>-1.80</td>
<td></td>
<td></td>
<td></td>
<td>-2.60</td>
<td>-0.38</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D-CCAPM</td>
<td></td>
<td>-0.62</td>
<td></td>
<td></td>
<td></td>
<td>-0.84</td>
<td>3.23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC-MY</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SIM</td>
<td></td>
<td>-1.79</td>
<td>-1.85</td>
<td></td>
<td></td>
<td>-1.88</td>
<td>0.36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-1.83</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-2.94</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1.80</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.73</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1.81</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.46</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.36</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-2.90</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-1.48</td>
</tr>
</tbody>
</table>
Table 8 (continued)
Quarterly analysis of some popular linear asset pricing models

Panel D: Model selection procedure using model misspecification-robust standard errors

<table>
<thead>
<tr>
<th>Model</th>
<th>vw</th>
<th>vw·ml</th>
<th>smb</th>
<th>hml</th>
<th>c</th>
<th>cn</th>
<th>cd</th>
<th>cay</th>
<th>c·cay</th>
<th>my</th>
<th>c·my</th>
<th>i_b</th>
<th>i_c</th>
<th>i_ac</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAPM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C-ML</td>
<td>1.51</td>
<td>−1.54</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FF3</td>
<td>−3.71</td>
<td>−0.28</td>
<td>−4.37</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCAPM</td>
<td>−1.27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC-CAY</td>
<td>−0.37</td>
<td></td>
<td>−1.60</td>
<td>−0.16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D-CCAPM</td>
<td>−1.64</td>
<td>−1.17</td>
<td>0.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC-MY</td>
<td>−1.74</td>
<td>−1.24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SIM</td>
<td>−2.93</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>−2.15</td>
<td>1.69</td>
<td>−1.10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>−2.26</td>
<td>1.77</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>−1.82</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The table presents the estimation results of eight asset pricing models. The models include the CAPM, the conditional CAPM (C-ML) of Santos and Veronesi (2006), the Fama and French (1993) three-factor model (FF3), the consumption CAPM (CCAPM), the conditional consumption CAPM (CC-CAY) of Lettau and Ludvigson (2001), the durable consumption CAPM (D-CCAPM) of Yogo (2006), the conditional consumption CAPM (CC-MY) proposed by Lustig and Van Nieuwerburgh (2005), and the sector investment model (SIM) of Li, Vassalou, and Xing (2006). The models are estimated using quarterly gross returns on the 25 Fama-French size and book-to-market ranked portfolios and the one-month T-bill. The data are from 1952 Q2 until 2007 Q4. Panel A reports the rank restriction test ($W^*$) and its p-value ($p$-val) of the null that $E[x_t(1, f_t)]$ has a column rank of one. In Panel B, we report the sample HJ-distance ($\hat{\delta}$), the Lagrange multiplier ($LM$) test, and the rank restriction test ($W^*$) with the corresponding p-values ($p$-val) for each model. The $t$-tests of the model selection procedures based on the standard errors under correct model specification and model misspecification are in Panels C and D, respectively. Each $t$-test is for the test of the null hypothesis that the coefficient associated with a given risk factor is equal to zero. We use boldface to highlight those cases in which the factors survive the model selection procedure at the 5% significance level.
Figure 1
Asymptotic distributions of $t_c(\hat{\gamma}_2)$ and $t_m(\hat{\gamma}_2)$ under misspecified models. The figure presents the probability density functions of the limiting distributions of $t_c(\hat{\gamma}_2)$ and $t_m(\hat{\gamma}_2)$, the $t$-statistics for the useless factor that use standard errors constructed under correctly specified and potentially misspecified models, respectively, for $N - K = 7$ (see part (b) of Theorem 2).
Figure 2
Limiting probabilities of rejection of the HJ-distance test. The figure presents the limiting probabilities of rejection of the HJ-distance test under correctly specified and misspecified models when one of the factors is useless.