ECONSTOR
Make Your Publications Visible.

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Gospodinov, Nikolay; Kan, Raymond; Robotti, Cesare

Working Paper

Robust iInference in linear asset pricing models

Working Paper, No. 2012-17

Provided in Cooperation with:
Federal Reserve Bank of Atlanta

Suggested Citation: Gospodinov, Nikolay; Kan, Raymond; Robotti, Cesare (2012) : Robust iInference in linear asset pricing models, Working Paper, No. 2012-17, Federal Reserve Bank of Atlanta, Atlanta, GA

This Version is available at:
https://hdl.handle.net/10419/70612

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

FEDERAL RESERVE BANK *of* ATLANTA **WORKING PAPER SERIES**

Robust Inference in Linear Asset Pricing Models

Nikolay Gospodinov, Raymond Kan, and Cesare Robotti

Working Paper 2012-17
November 2012

Abstract: We derive new results on the asymptotic behavior of the estimated parameters of a linear asset pricing model and their associated *t*-statistics in the presence of a factor that is independent of the returns. The inclusion of this "useless" factor in the model leads to a violation of the full rank (identification) condition and renders the inference nonstandard. We show that the estimated parameter associated with the useless factor diverges with the sample size but the misspecification-robust *t*-statistic is still well-behaved and has a standard normal limiting distribution. The asymptotic distributions of the estimates of the remaining parameters and the model specification test are also affected by the presence of a useless factor and are nonstandard. We propose a robust and easy-to-implement model selection procedure that restores the standard inference on the parameters of interest by identifying and removing the factors that do not contribute to improved pricing. The finite-sample properties of our asymptotic approximations and the practical relevance of our results are illustrated using simulations and an empirical application.

JEL classification: G12, C13, C32

Key words: asset pricing models, Hansen-Jagannathan distance, model selection, model misspecification

---

Gospodinov gratefully acknowledges financial support from Fonds de recherche sur la société et la culture (FQRSC), Institut de Finance Mathematique de Montreal (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. They also thank seminar participants at the Federal Reserve Bank of Atlanta, Singapore Management University, the University of Georgia, and participants at the third annual CIRPEE Applied Financial Time Series Workshop at HEC (Montreal) and Mathematical Finance Days 2012 for helpful comments. The views expressed here are the authors' and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System. Any remaining errors are the authors' responsibility.

Please address questions regarding content to Nikolay Gospodinov, Department of Economics, Concordia University, 1455 de Maisonneuve Boulevard West, Montreal, Quebec H3G 1M8, Canada, and CIREQ, 514-848-2424, 514-848-4536 (fax), nikolay.gospodinov@concordia.ca; Raymond Kan, Joseph L. Rotman School of Management, University of Toronto, 105 St. George Street, Toronto, Ontario M5S 3E6, Canada, kan@chass.utoronto.ca; or Cesare Robotti, Research Department, Federal Reserve Bank of Atlanta, 1000 Peachtree Street N.E., Atlanta, GA 30309, and EDHEC Risk Institute, cesare.robotti@atl.frb.org.

Federal Reserve Bank of Atlanta working papers, including revised versions, are available on the Atlanta Fed's website at frbatlanta.org/pubs/WP/. Use the WebScriber Service at frbatlanta.org to receive e-mail notifications about new papers.

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 (2012b) 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 asset pricing models that are potentially misspecified, 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 Wald test for the useless factor is asymptotically distributed as a chi-squared random variable with

[1]While we study explicitly only the GMM estimator based on the HJ-distance, our results continue to hold for the class of optimal GMM estimators.

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 estimate 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 that do not provide improved pricing. 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 (ICAPM) used by Petkova (2006). We also study the well-known "three-factor model" of Fama and

[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.

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 rank restriction test proposed by Gospodinov, Kan, and Robotti (2012a) 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 and book-to-market 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 I reviews some of the main results for asymptotically valid inference under potential model misspecification. In Section II, 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 III reports results from a Monte Carlo simulation experiment. In Section IV, we conduct an empirical investigation of some popular asset pricing models with traded and non-traded factors. Section V concludes.

I. Asymptotic Inference with Useful Factors

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

$$y_t(\gamma_1) = \tilde{f}_t'\gamma_1 \tag{1}$$

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 γ_1 is a K-vector of SDF parameters with generic element γ_{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. Note that our setup can accommodate conditional asset pricing models that parameterize γ_1 as a linear function of a vector of conditioning variables z_{t-1}. In this case, $\tilde{f}_t$ in (1) effectively denotes the vector of scaled factors $\tilde{f}_t \otimes z_{t-1}$, where $\otimes$ is the Kronecker product.

Define the model pricing errors as

$$e(\gamma_1) = E[x_t \tilde{f}_t' \gamma_1 - q] = B\gamma_1 - q. \tag{2}$$

If there exists no value of γ_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 γ_1^* is defined as the solution to the quadratic minimization problem

$$\gamma_1^* = \underset{\gamma_1 \in \Gamma_1}{\arg\min}\ e(\gamma_1)' W e(\gamma_1) \tag{3}$$

for some symmetric and positive-definite weighting matrix W, where Γ_1 denotes the parameter space.

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

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

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 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 $\tilde{\gamma}_1$ of γ_1^* is obtained by minimizing the sample analog of (3):

$$\tilde{\gamma}_1 = \underset{\gamma_1 \in \Gamma_1}{\arg\min}\ \hat{e}(\gamma_1)' \hat{U}^{-1} \hat{e}(\gamma_1), \tag{4}$$

[3] When $q = 0_N$, 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).

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

$$\hat{B} = \frac{1}{T}\sum_{t=1}^{T} x_t \tilde{f}_t'. \tag{5}$$

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

$$\tilde{\gamma}_1 = (\hat{B}'\hat{U}^{-1}\hat{B})^{-1}(\hat{B}'\hat{U}^{-1}q). \tag{6}$$

For ease of exposition, we assume that $e_t(\gamma_1^*) = x_t\tilde{f}_t'\gamma_1^* - q$ forms a stationary and ergodic martingale difference sequence and define $S = E[e_t(\gamma_1^*)e_t(\gamma_1^*)']$. Under suitable regularity conditions, Kan and Robotti (2009) show that

$$\sqrt{T}(\tilde{\gamma}_1 - \gamma_1^*) \xrightarrow{d} N(0_K, \Sigma_{\tilde{\gamma}_1}), \tag{7}$$

where $\Sigma_{\tilde{\gamma}_1} = E[h_t h_t']$,

$$h_t = (B'U^{-1}B)^{-1}B'U^{-1}e_t(\gamma_1^*) + (B'U^{-1}B)^{-1}(\tilde{f}_t - B'U^{-1}x_t)u_t \tag{8}$$

and

$$u_t = e(\gamma_1^*)'U^{-1}x_t. \tag{9}$$

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

$$h_t^0 = (B'U^{-1}B)^{-1}B'U^{-1}e_t(\gamma_1^*) \tag{10}$$

and the asymptotic covariance matrix of $\sqrt{T}(\tilde{\gamma}_1 - \gamma_1^*)$ is simplified to

$$\Sigma_{\tilde{\gamma}_1}^0 = E[h_t^0 h_t^{0\prime}] = (B'U^{-1}B)^{-1}B'U^{-1}SU^{-1}B(B'U^{-1}B)^{-1}. \tag{11}$$

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 $\boldsymbol{\iota}_i$ with one for its i-th element and zero otherwise (the length of $\boldsymbol{\iota}_i$ 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

$$t_m(\tilde{\gamma}_{1i}) = \frac{\tilde{\gamma}_{1i} - \gamma_{1i}^*}{\sqrt{\boldsymbol{\iota}_i'\hat{\Sigma}_{\tilde{\gamma}_1}\boldsymbol{\iota}_i}} \xrightarrow{d} N(0,1), \tag{12}$$

where $\hat{\Sigma}_{\tilde{\gamma}_1}$ is a consistent estimator of $\Sigma_{\tilde{\gamma}_1}$. 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

$$t_c(\tilde{\gamma}_{1i}) = \frac{\tilde{\gamma}_{1i} - \gamma^*_{1i}}{\sqrt{\boldsymbol{\iota}_i' \hat{\Sigma}^0_{\tilde{\gamma}_1} \boldsymbol{\iota}_i}}, \tag{13}$$

where $\hat{\Sigma}^0_{\tilde{\gamma}_1}$ is a consistent estimator of $\Sigma^0_{\tilde{\gamma}_1}$. If the model is indeed correctly specified, the t-test $t_c(\tilde{\gamma}_{1i})$ is asymptotically distributed as a standard normal random variable

$$t_c(\tilde{\gamma}_{1i}) \xrightarrow{d} N(0,1). \tag{14}$$

However, using the result in (7)–(8), we have that under misspecified models

$$t_c(\tilde{\gamma}_{1i}) \xrightarrow{d} N\left(0, \frac{\boldsymbol{\iota}_i' \Sigma_{\tilde{\gamma}_1} \boldsymbol{\iota}_i}{\boldsymbol{\iota}_i' \Sigma^0_{\tilde{\gamma}_1} \boldsymbol{\iota}_i}\right). \tag{15}$$

Furthermore, under the assumption that x_t and f_t are multivariate elliptically distributed, it can be shown (Kan and Robotti, 2009) that $(\boldsymbol{\iota}_i' \Sigma_{\tilde{\gamma}_1} \boldsymbol{\iota}_i)/(\boldsymbol{\iota}_i' \Sigma^0_{\tilde{\gamma}_1} \boldsymbol{\iota}_i) > 1$, 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 (8) which is used for computing the covariance matrix $\Sigma_{\tilde{\gamma}_1}$ under misspecification. It proves useful to rewrite h_t as

$$h_t = h^0_t + (B' U^{-1} B)^{-1} (\tilde{f}_t - B' U^{-1} x_t) u_t. \tag{16}$$

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.

II. 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

$$y_t = \tilde{f}_t'\gamma_1 + g_t\gamma_2, \tag{17}$$

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

$$d = E[x_t g_t] = 0_N \tag{18}$$

and

$$E[x_t x_t' g_t^2] = E[E[x_t x_t'|g_t]g_t^2] = UE[g_t^2] = U. \tag{19}$$

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

$$e(\gamma) = B\gamma_1 + d\gamma_2 - q = B\gamma_1 - q \tag{20}$$

is independent of the choice of γ_2. For the pseudo-true values of the SDF parameters, we can set γ_1^* as in (3) but the parameters associated with the useless factor (γ_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.

While the pseudo-true values of γ_2^* are not identified, the sample estimates of the SDF parameters are always identified and they are given by

$$\hat{\gamma} = (\hat{D}'\hat{U}^{-1}\hat{D})^{-1}(\hat{D}'\hat{U}^{-1}q). \tag{21}$$

Note that this is equivalent to running an ordinary least squares (OLS) regression of $\hat{U}^{-\frac{1}{2}}q$ on $\hat{U}^{-\frac{1}{2}}\hat{B}$ and $\hat{U}^{-\frac{1}{2}}\hat{d}$. In order to obtain $\hat{\gamma}_2$, we can project $\hat{U}^{-\frac{1}{2}}q$ and $\hat{U}^{-\frac{1}{2}}\hat{d}$ on $\hat{U}^{-\frac{1}{2}}\hat{B}$, and then

[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.

regress the residuals from the first projection onto the residuals from the second projection. It follows that

$$\hat{\gamma}_2 = \frac{\hat{d}'\hat{U}^{-\frac{1}{2}}\left[I_N - \hat{U}^{-\frac{1}{2}}\hat{B}(\hat{B}'\hat{U}^{-1}\hat{B})^{-1}\hat{B}'\hat{U}^{-\frac{1}{2}}\right]\hat{U}^{-\frac{1}{2}}q}{\hat{d}'\hat{U}^{-\frac{1}{2}}\left[I_N - \hat{U}^{-\frac{1}{2}}\hat{B}(\hat{B}'\hat{U}^{-1}\hat{B})^{-1}\hat{B}'\hat{U}^{-\frac{1}{2}}\right]\hat{U}^{-\frac{1}{2}}\hat{d}}. \tag{22}$$

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

$$\begin{aligned}\hat{\gamma}_1 &= (\hat{B}'\hat{U}^{-\frac{1}{2}}[I_N - \hat{U}^{-\frac{1}{2}}\hat{d}(\hat{d}'\hat{U}^{-\frac{1}{2}}\hat{d})^{-1}\hat{d}'\hat{U}^{-\frac{1}{2}}]\hat{U}^{-\frac{1}{2}}\hat{B})^{-1} \\ &\quad \times \hat{B}'\hat{U}^{-\frac{1}{2}}[I_N - \hat{U}^{-\frac{1}{2}}\hat{d}(\hat{d}'\hat{U}^{-\frac{1}{2}}\hat{d})^{-1}\hat{d}'\hat{U}^{-\frac{1}{2}}]\hat{U}^{-\frac{1}{2}}q. \end{aligned} \tag{23}$$

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 $\tilde{B} = U^{-\frac{1}{2}}B$, $\tilde{q} = U^{-\frac{1}{2}}q$, $S = E[e_t(\gamma^*)e_t(\gamma^*)']$, and P be an $N \times (N-K)$ orthonormal matrix whose columns are orthogonal to $\tilde{B}$ so that $PP' = I_N - \tilde{B}(\tilde{B}'\tilde{B})^{-1}\tilde{B}'$. Also, let $z \sim N(0_N, I_N)$ and $y \sim N(0_N, U^{-\frac{1}{2}}SU^{-\frac{1}{2}})$, and they are independent of each other. Finally, we define $w = P'z \sim N(0_{N-K}, I_{N-K})$, $s = (\tilde{q}'Pw)/(\tilde{q}'PP'\tilde{q})^{\frac{1}{2}} \sim N(0,1)$, $u = P'y \sim N(0_{N-K}, V_u)$ with $V_u = P'U^{-\frac{1}{2}}SU^{-\frac{1}{2}}P$, and $r = (\tilde{B}'\tilde{B})^{-\frac{1}{2}}\tilde{B}'y \sim N(0_K, V_r)$ with $V_r = (\tilde{B}'\tilde{B})^{-\frac{1}{2}}\tilde{B}'U^{-\frac{1}{2}}SU^{-\frac{1}{2}}\tilde{B}(\tilde{B}'\tilde{B})^{-\frac{1}{2}}$.

Theorem 1. *Assume that $N > K+1$, $[x_t',\ f_t',\ g_t]'$ are jointly stationary and ergodic processes and $e_t(\gamma^*) = x_t\tilde{f}_t'\gamma_1^* - q$ forms a martingale difference sequence with $E_{t-1}[e_t(\gamma^*)] = 0_N$ (a.s.), $E_{t-1}[e_t(\gamma^*)e_t(\gamma^*)'] = S$ (a.s.) and $\sup_t E\,\|e_t(\gamma^*)\|^4 < \infty$, where S is a positive definite matrix, $E_{t-1}[\cdot]$ denotes the expectation conditional on $\{e_{t-1}(\gamma^*), e_{t-2}(\gamma^*), \ldots\}$, and $\|\cdot\|$ is the Euclidean norm.*

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

$$\sqrt{T}(\hat{\gamma}_1 - \gamma_1^*) \xrightarrow{d} (\tilde{B}'\tilde{B})^{-\frac{1}{2}}\left[r - \frac{w'u}{w'w}(\tilde{B}'\tilde{B})^{-\frac{1}{2}}\tilde{B}'z\right], \tag{24}$$

and

$$\hat{\gamma}_2 \xrightarrow{d} \frac{w'u}{w'w}. \tag{25}$$

(b) *If $\delta > 0$, i.e., the model is misspecified, we have*

$$\hat{\gamma}_1 - \gamma_1^* \xrightarrow{d} -\frac{\delta s}{w'w}(\tilde{B}'\tilde{B})^{-1}\tilde{B}'z, \tag{26}$$

and

$$\frac{1}{\sqrt{T}}\hat{\gamma}_2 \xrightarrow{d} \frac{\delta s}{w'w}. \tag{27}$$

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.[5] 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 $T^{\frac{1}{2}}$ 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 $\hat{\Sigma}^0_{\hat{\gamma}} = \frac{1}{T}\sum_{t=1}^T \hat{h}^0_t \hat{h}^{0\prime}_t$ and $\hat{\Sigma}_{\hat{\gamma}} = \frac{1}{T}\sum_{t=1}^T \hat{h}_t \hat{h}'_t$, where

$$\hat{h}^0_t = (\hat{D}'\hat{U}^{-1}\hat{D})^{-1}\hat{D}'\hat{U}^{-1}\hat{e}_t, \tag{28}$$

$$\hat{h}_t = \hat{h}^0_t + (\hat{D}'\hat{U}^{-1}\hat{D})^{-1}([\tilde{f}'_t,\ g_t]' - \hat{D}'\hat{U}^{-1}x_t)\hat{e}'\hat{U}^{-1}x_t, \tag{29}$$

and $\hat{e}_t = x_t(\tilde{f}'_t\hat{\gamma}_1 + g_t\hat{\gamma}_2) - q$. 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

$$t_c(\hat{\gamma}_{1i}) = \frac{\sqrt{T}(\hat{\gamma}_{1i} - \gamma^*_{1i})}{\sqrt{\boldsymbol{\iota}'_i\hat{\Sigma}^0_{\hat{\gamma}}\boldsymbol{\iota}_i}} \tag{30}$$

and

$$t_c(\hat{\gamma}_2) = \frac{\sqrt{T}\hat{\gamma}_2}{\sqrt{\boldsymbol{\iota}'_{K+1}\hat{\Sigma}^0_{\hat{\gamma}}\boldsymbol{\iota}_{K+1}}}. \tag{31}$$

[5]The limiting random variable has mean zero and variance $\mathrm{tr}(V_u)/[(N-K)(N-K-2)]$, where $\mathrm{tr}(\cdot)$ is the trace operator.

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

$$t_m(\hat{\gamma}_{1i}) = \frac{\sqrt{T}(\hat{\gamma}_{1i} - \gamma_{1i}^*)}{\sqrt{\boldsymbol{\iota}_i' \hat{\Sigma}_{\hat{\gamma}} \boldsymbol{\iota}_i}} \tag{32}$$

and

$$t_m(\hat{\gamma}_2) = \frac{\sqrt{T}\hat{\gamma}_2}{\sqrt{\boldsymbol{\iota}_{K+1}' \hat{\Sigma}_{\hat{\gamma}} \boldsymbol{\iota}_{K+1}}}. \tag{33}$$

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

$$\hat{e}_t = (T^{-\frac{1}{2}}\hat{\gamma}_2)(T^{\frac{1}{2}} x_t g_t) + O_p(1) \tag{34}$$

and

$$\frac{\hat{S}}{T} = (T^{-\frac{1}{2}}\hat{\gamma}_2)^2 U + o_p(1), \tag{35}$$

so $\hat{S}$ diverges at rate T. In contrast, for correctly specified models, we have

$$\hat{S} = S + \hat{\gamma}_2^2 U + o_p(1), \tag{36}$$

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* $\epsilon_t = x_t - B(E[\tilde{f}_t \tilde{f}_t'])^{-1}\tilde{f}_t$.[6] *If* $\delta = 0$, *i.e., the model is*

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

correctly specified, we have

$$t_c(\hat{\gamma}_{1i}) \xrightarrow{d} \frac{\tilde{u}\tilde{z}_i + \sqrt{\lambda_i}\sqrt{w'w}\tilde{r}_i}{\left[\lambda_i w'w + \tilde{z}_i^2 + \tilde{u}^2\left(1 + \frac{\tilde{z}_i^2}{w'w}\right)\right]^{\frac{1}{2}}}, \tag{37}$$

$$t_m(\hat{\gamma}_{1i}) \xrightarrow{d} \frac{\tilde{u}\tilde{z}_i + \sqrt{\lambda_i}\sqrt{w'w}\tilde{r}_i}{\left[\lambda_i w'w + \tilde{z}_i^2 + \tilde{u}^2\left(1 + \frac{\tilde{z}_i^2}{w'w}\right) + \frac{\tilde{z}_i^2 v}{w'w}\right]^{\frac{1}{2}}}, \tag{38}$$

$$t_c(\hat{\gamma}_2) \xrightarrow{d} \frac{\tilde{u}}{\left(1 + \frac{\tilde{u}^2}{w'w}\right)^{\frac{1}{2}}}, \tag{39}$$

$$t_m(\hat{\gamma}_2) \xrightarrow{d} \frac{\tilde{u}}{\left(1 + \frac{\tilde{u}^2 + v}{w'w}\right)^{\frac{1}{2}}}, \tag{40}$$

where λ_i is a positive constant and its explicit expression is given in the Appendix.

(b) *Suppose that the assumptions in Theorem 1 hold and denote the sign operator by* $\operatorname{sgn}(\cdot)$. *If* $\delta > 0$, *i.e., the model is misspecified, we have*

$$t_c(\hat{\gamma}_{1i}) \xrightarrow{d} \frac{\tilde{z}_i}{\left(1 + \frac{\tilde{z}_i^2}{w'w}\right)^{\frac{1}{2}}}, \tag{41}$$

$$t_m(\hat{\gamma}_{1i}) \xrightarrow{d} N\left(0, \frac{1}{4}\right), \tag{42}$$

$$t_c(\hat{\gamma}_2) \xrightarrow{d} \operatorname{sgn}(s)\sqrt{w'w}, \tag{43}$$

$$t_m(\hat{\gamma}_2) \xrightarrow{d} N(0,1). \tag{44}$$

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* χ_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* χ_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 γ 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. 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.

Figure 1 about here

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 α 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 I.

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

$$\hat{\delta}^2 = \hat{e}'\hat{U}^{-1}\hat{e}. \tag{45}$$

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

$$T\hat{\delta}^2 \xrightarrow{d} \sum_{i=1}^{N-K} \xi_i X_i, \tag{46}$$

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

$$S^{\frac{1}{2}}U^{-1}S^{\frac{1}{2}} - S^{\frac{1}{2}}U^{-1}B(B'U^{-1}B)^{-1}B'U^{-1}S^{\frac{1}{2}}. \tag{47}$$

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

$$\hat{S}^{\frac{1}{2}}\hat{U}^{-1}\hat{S}^{\frac{1}{2}} - \hat{S}^{\frac{1}{2}}\hat{U}^{-1}\hat{B}(\hat{B}'\hat{U}^{-1}\hat{B})^{-1}\hat{B}'\hat{U}^{-1}\hat{S}^{\frac{1}{2}}. \tag{48}$$

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_α *be the* $100(1-\alpha)$-th *percentile of* χ^2_{N-K-1}.

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

$$T\hat{\delta}^2 \xrightarrow{d} E[(\tilde{f}_t'\gamma_1^*)^2]\chi^2_{N-K-1} \tag{49}$$

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

$$\int_0^1 P\left[\chi^2_{N-K-1} > \frac{c_\alpha}{q}\right] f_{Q_1}(q)\mathrm{d}q < \alpha. \tag{50}$$

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

$$\hat{\delta}^2 \xrightarrow{d} \delta^2 Q_2 \tag{51}$$

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

$$\int_0^1 P\left[\chi^2_{N-K} > \frac{c_\alpha q}{1-q}\right] f_{Q_2}(q)\mathrm{d}q < 1. \tag{52}$$

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. Hence, the test will not reject a misspecified model with a probability approaching one when a useless factor is present.

Note that the limiting probabilities of rejection in (50) and (52) are only functions of the significance level α and the degree of over-identification $N-K$. Figure 2 plots these probabilities for different significance levels ($\alpha = 0.01,\ 0.05$, 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.

Figure 2 about here

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-tests is necessary. Specifically, in the first stage, we estimate the full model and the factors for which the null of zero risk premium is not rejected are eliminated from the model. The model is then re-estimated with only the factors that survive the first stage at the pre-specified nominal level. This procedure is repeated until 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-tests. The appeal of this model selection procedure is that the inference for testing the significance of the SDF parameters γ is standard (i.e., it 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.

III. 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.

A. 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 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).[7] Similarly, the mean and variance of the simulated useful factor are calibrated to the 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, 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 I to III, 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 γ_i^*'s for the constant and the useful factor are the chosen pseudo SDF parameters, and the γ_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.

Table I about here

The results in Table I.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 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 II and III present the empirical size of the t-tests in the presence of a useless factor.

Tables II and III about here

[7] See Gospodinov, Kan, and Robotti (2012b) for a detailed description of how the parameters are chosen in the different simulation designs.

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 II 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 for significance of 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.

B. 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.

Table IV about here

The results in Table IV 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.

C. 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 V reports the survival rates of different factors when using the sequential procedure described at the end of Section II. 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 I–IV, 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 V 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 provide improved pricing. 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 47% of the times for $T = 1000$. In contrast, the procedure based on t_m will retain the useless factor only about 4% of the times 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.[8]

IV. 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 pricing 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.[9]

A. 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.

A.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

$$y_t^{CAPM}(\gamma) = \gamma_0 + \gamma_1 vw_t,$$

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

[8]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.

[9]Some additional empirical results are provided in an online appendix available on the authors' websites.

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.[10] The SDF specification is

$$y_t^{C-LAB}(\gamma) = \gamma_0 + \gamma_1 vw_t + \gamma_2 prem_{t-1} + \gamma_3 lab_t,$$

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

$$y_t^{FF3}(\gamma) = \gamma_0 + \gamma_1 vw_t + \gamma_2 smb_t + \gamma_3 hml_t,$$

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 the 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*.

[10] All bond yield data are from this source unless noted otherwise.

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.[11] This specification is obtained by scaling the constant term and the c factor of a linearized consumption CAPM by a constant and 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

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

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.

A.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 $K-1$. Under the null hypothesis H_0 : rank$(B) = K - 1$, there exists a nonzero K-vector $\tilde{c} = [-1,\ c']$

[11] Following Jørgensen and Attanasio (2003), we linearly interpolate the quarterly values of *cay* to permit analysis at the monthly frequency.

such that $B\tilde{c} = 0_N$. Let $\hat{M}$ be a consistent estimator of the asymptotic covariance matrix of $\sqrt{T}\text{vec}(\hat{B} - B)$, where $\text{vec}(\cdot)$ is the vec operator, and $Q(c) = [-1,\ c'] \otimes I_N$. Consider the rank restriction test

$$\mathcal{W}^* = \min_c T\text{vec}(\hat{B})'Q(c)'[Q(c)\hat{M}Q(c)']^{-1}Q(c)\text{vec}(\hat{B}). \tag{53}$$

Under some regularity conditions and $H_0 : \text{rank}(B) = K - 1$ (see Cragg and Donald, 1997, and Gospodinov, Kan, and Robotti, 2012a),

$$\mathcal{W}^* \xrightarrow{d} \chi^2_{N-(K-1)}. \tag{54}$$

Table VI.A reports the rank restriction test ($\mathcal{W}^*$) and its p-value (p-val) of the null that $E[x_t(1,\ f_{it})]$ 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 (2012b) 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 VI.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 often misspecified and very poorly identified.

Table VI about here

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 VI present the t-tests under correct model specification and potential model misspecification. The t-tests under correct model specification suggest that the SDF parameters on lab in C-LAB, $term$ and def in ICAPM, c in CCAPM, c and cay in CC-CAY, and c_{nd} in D-CCAPM are statistically significant at the 5% nominal 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. In contrast, using the misspecification-robust t-tests, we do not reject the null hypothesis of a zero SDF parameter for these variables.

Finally, Panel D shows that only vw in CAPM and FF3 and hml in FF3 survive the first stage of the sequential procedure based on misspecification-robust t-tests at the 5% significance level. Since smb is not found to be statistically significant in the first stage, we drop this factor from the analysis and estimate an asset pricing specification with vw and hml as the only two risk factors. We still find strong evidence of pricing for the vw and hml factors with misspecification-robust t-statistics of -4.20 and -4.89, respectively.

B. 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 vw and vw scaled by the labor income-consumption ratio (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 (my), c, and the interaction term $my \cdot c$ 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 VII, with Panel A showing that for all factors except for vw, $vw{\cdot}ml$, smb, 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.[12] This clearly points to the need of statistical procedures that are robust to model misspecification and weak identification.

Table VII about here

Panel D of Table VII shows that all the non-traded factors except for i_h do not survive the first stage of the model selection procedure based on the misspecification-robust t-test.[13] This stands in sharp contrast to the results in Panel C of Table VII where the t-ratio under correctly specified models is employed. However, our theoretical and simulation analyses clearly showed that relying on the t-ratio 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 C-ML. The t-statistic under correctly specified models on the risk premium estimator for the $vw{\cdot}ml$ factor goes from -2.10 to -1.54 when misspecification and weak identification are taken into account.

Finally, as in the monthly case, only 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.

V. 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

[12] Using the LM test and a 5% significance level, we can reject the null of correct specification for all models.

[13] Since i_h is the only factor in SIM that passes the first stage of the proposed sequential procedure at the 5% nominal level (misspecification-robust t-statistic of -2.15), we drop i_c and i_{nc} from the analysis and estimate an asset pricing specification with i_h as the only risk factor. The misspecification-robust t-statistic of -1.82, however, indicates that i_h is not a priced factor in the second stage of our model selection procedure at the 5% significance level.

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 asymptotically 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 investigate the empirical performance of several prominent asset pricing models with traded and non-traded factors. We find that only the market factor and the book-to-market factor of Fama and French (1993) survive the model selection procedure based on misspecification-robust t-ratios. For non-traded factors, there is not enough statistical evidence for us to conclude that they are important in explaining the cross-sectional differences in expected asset returns.

Appendix: Preliminary Lemma and Proofs of Main Results

A.1 Preliminary Lemma

Lemma A.1. *Let*

$$x_t = BS_{\tilde{f}}^{-1}\tilde{f}_t + \epsilon_t, \tag{A.1}$$

where $B = E[x_t\tilde{f}_t']$, $S_{\tilde{f}} = E[\tilde{f}_t\tilde{f}_t']$ *and* $E[\epsilon_t|\tilde{f}_t] = 0_N$. *Suppose* $\mathrm{Cov}[\epsilon_t\epsilon_t', (\tilde{f}_t'\gamma_1^*)^2] = 0_{N\times N}$ *(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*

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

where $U = E[x_tx_t']$ *and* C *is a symmetric* $(K+1)\times(K+1)$ *matrix.*

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

$$S = E[x_tx_t'(\tilde{f}_t'\gamma_1^*)^2] - qq' = E[x_tx_t'(\tilde{f}_t'\gamma_1^*)^2] - B\gamma_1^*\gamma_1^*B'. \tag{A.3}$$

For the first term, we have

$$\begin{aligned} E[x_tx_t'(\tilde{f}_t'\gamma_1^*)^2] &= E[x_tx_t']E[(\tilde{f}_t'\gamma_1^*)^2] + \mathrm{Cov}[x_tx_t', (\tilde{f}_t'\gamma_1^*)^2] \\ &= E[(\tilde{f}_t'\gamma_1^*)^2]U + \mathrm{Cov}[BS_{\tilde{f}}^{-1}\tilde{f}_t\tilde{f}_t'S_{\tilde{f}}^{-1}B' + \epsilon_t\epsilon_t', (\tilde{f}_t'\gamma_1^*)^2] \\ &= E[(\tilde{f}_t'\gamma_1^*)^2]U + BS_{\tilde{f}}^{-1}\mathrm{Cov}[\tilde{f}_t\tilde{f}_t', (\tilde{f}_t'\gamma_1^*)^2]S_{\tilde{f}}^{-1}B', \end{aligned} \tag{A.4}$$

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

$$S = E[(\tilde{f}_t'\gamma_1^*)^2]U + BCB', \tag{A.5}$$

where

$$C = S_{\tilde{f}}^{-1}\mathrm{Cov}[\tilde{f}_t\tilde{f}_t', (\tilde{f}_t'\gamma_1^*)^2]S_{\tilde{f}}^{-1} - \gamma_1^*\gamma_1^{*\prime}. \tag{A.6}$$

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

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

and

$$-\sqrt{T}\hat{U}^{-\frac{1}{2}}(\hat{B} - B)\gamma_1^* \xrightarrow{d} y \sim N(0_N, (\gamma_1^{*\prime} \otimes U^{-\frac{1}{2}})V_B(\gamma_1^* \otimes U^{-\frac{1}{2}})), \tag{A.8}$$

where V_B is the asymptotic covariance of $\sqrt{T}\text{vec}(\hat{B} - B)$ and $\text{vec}(\cdot)$ is the vec operator. In addition, y and z are independent of each other. Since

$$V_B = E[\text{vec}(x_t\tilde{f}_t' - B)\text{vec}(x_t\tilde{f}_t' - B)'], \tag{A.9}$$

it follows that the covariance matrix of y is given by

$$V_y = (\gamma_1^{*\prime} \otimes U^{-\frac{1}{2}})V_B(\gamma_1^* \otimes U^{-\frac{1}{2}}) = E[m_t m_t'], \tag{A.10}$$

where

$$m_t = U^{-\frac{1}{2}}(x_t\tilde{f}_t' - B)\gamma_1^* = U^{-\frac{1}{2}}(x_t\tilde{f}_t'\gamma_1^* - q) = U^{-\frac{1}{2}}e_t(\gamma^*), \tag{A.11}$$

and the second equality is obtained by using the fact that $B\gamma_1^* = q$ because the model is correctly specified. Therefore, we have $V_y = U^{-\frac{1}{2}}SU^{-\frac{1}{2}}$ for correctly specified models.

Then, using that $q = B\gamma_1^*$, we can write (23) as

$$\begin{aligned}
\sqrt{T}(\hat{\gamma}_1 - \gamma_1^*) &= (\hat{B}'\hat{U}^{-\frac{1}{2}}[I_N - \hat{U}^{-\frac{1}{2}}\hat{d}(\hat{d}'\hat{U}^{-\frac{1}{2}}\hat{d})^{-1}\hat{d}'\hat{U}^{-\frac{1}{2}}]\hat{U}^{-\frac{1}{2}}\hat{B})^{-1} \\
&\quad \times \hat{B}'\hat{U}^{-\frac{1}{2}}[I_N - \hat{U}^{-\frac{1}{2}}\hat{d}(\hat{d}'\hat{U}^{-\frac{1}{2}}\hat{d})^{-1}\hat{d}'\hat{U}^{-\frac{1}{2}}]\sqrt{T}\hat{U}^{-\frac{1}{2}}(B - \hat{B})\gamma_1^* \\
&\xrightarrow{d} (\tilde{B}'[I_N - z(z'z)^{-1}z']\tilde{B})^{-1}\tilde{B}'[I_N - z(z'z)^{-1}z']y \\
&= (\tilde{B}'[I_N - z(z'z)^{-1}z']\tilde{B})^{-1}\tilde{B}'[I_N - z(z'z)^{-1}z'][PP' + \tilde{B}(\tilde{B}'\tilde{B})^{-1}\tilde{B}']y \\
&= -(\tilde{B}'[I_N - z(z'z)^{-1}z']\tilde{B})^{-1}\frac{\tilde{B}'zz'PP'y}{z'z} + (\tilde{B}'\tilde{B})^{-1}\tilde{B}'y.
\end{aligned} \tag{A.12}$$

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

$$(\tilde{B}'[I_N - z(z'z)^{-1}z']\tilde{B})^{-1} = (\tilde{B}'\tilde{B})^{-1} + \frac{(\tilde{B}'\tilde{B})^{-1}\tilde{B}'zz'\tilde{B}(\tilde{B}'\tilde{B})^{-1}}{w'w} \tag{A.13}$$

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

$$\sqrt{T}(\hat{\gamma}_1 - \gamma_1^*) \xrightarrow{d} (\tilde{B}'\tilde{B})^{-\frac{1}{2}}\left[-\frac{w'u}{w'w}(\tilde{B}'\tilde{B})^{-\frac{1}{2}}\tilde{B}'z + r\right]. \tag{A.14}$$

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 $\hat{M} = I_N - \hat{U}^{-\frac{1}{2}}\hat{B}(\hat{B}'\hat{U}^{-1}\hat{B})^{-1}\hat{B}'\hat{U}^{-\frac{1}{2}}$. Using that $q = B\gamma_1^*$ and $\hat{M}\hat{U}^{-\frac{1}{2}}\hat{B} = 0_{N\times K}$, we obtain

$$\sqrt{T}\hat{M}\hat{U}^{-\frac{1}{2}}q = \sqrt{T}\hat{M}\hat{U}^{-\frac{1}{2}}(B-\hat{B})\gamma_1^* \xrightarrow{d} My, \tag{A.15}$$

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

$$\hat{\gamma}_2 = \frac{(\sqrt{T}\hat{U}^{-\frac{1}{2}}\hat{d})'(\sqrt{T}\hat{M}\hat{U}^{-\frac{1}{2}}(B-\hat{B})\gamma_1^*)}{(\sqrt{T}\hat{U}^{-\frac{1}{2}}\hat{d})'\hat{M}(\sqrt{T}\hat{U}^{-\frac{1}{2}}\hat{d})}. \tag{A.16}$$

Then, from (A.7), (A.8) and $\hat{M} \xrightarrow{p} M = PP'$, we get

$$\hat{\gamma}_2 \xrightarrow{d} \frac{z'My}{z'Mz} = \frac{(P'z)'(P'y)}{(P'z)'(P'z)} = \frac{w'u}{w'w}. \tag{A.17}$$

This completes the proof of part (a) of Theorem 1.

part (b): Using the fact that $\hat{U}^{-\frac{1}{2}}\hat{B} \xrightarrow{\text{a.s.}} \tilde{B}$ and $\sqrt{T}\hat{U}^{-\frac{1}{2}}\hat{d} \xrightarrow{d} z$, we can obtain the limiting distribution of $\hat{\gamma}_1$ in (23) as

$$\hat{\gamma}_1 \xrightarrow{d} (\tilde{B}'[I_N - z(z'z)^{-1}z']\tilde{B})^{-1}\tilde{B}'[I_N - z(z'z)^{-1}z']\tilde{q}. \tag{A.18}$$

Using (A.13) and the fact that $\gamma_1^* = (\tilde{B}'\tilde{B})^{-1}\tilde{B}'\tilde{q}$, we obtain

$$\begin{aligned}
\hat{\gamma}_1 - \gamma_1^* \xrightarrow{d} & \left[(\tilde{B}'\tilde{B})^{-1} + \frac{(\tilde{B}'\tilde{B})^{-1}\tilde{B}'zz'\tilde{B}(\tilde{B}'\tilde{B})^{-1}}{w'w}\right]\left(\tilde{B}'q - \frac{\tilde{B}'zz'\tilde{q}}{z'z}\right) - (\tilde{B}'\tilde{B})^{-1}\tilde{B}'\tilde{q} \\
= & -(\tilde{B}'\tilde{B})^{-1}\tilde{B}'z\frac{z'\tilde{q}}{z'z} + (\tilde{B}'\tilde{B})^{-1}\tilde{B}'z\frac{z'\tilde{B}(\tilde{B}'\tilde{B})^{-1}\tilde{B}'\tilde{q}}{w'w} - (\tilde{B}'\tilde{B})^{-1}\tilde{B}'z\frac{z'\tilde{q}}{z'z}\frac{z'\tilde{B}(\tilde{B}'\tilde{B})^{-1}\tilde{B}'z}{w'w} \\
= & -(\tilde{B}'\tilde{B})^{-1}\tilde{B}'z\frac{z'\tilde{q}}{w'w} + (\tilde{B}'\tilde{B})^{-1}\tilde{B}'z\frac{z'\tilde{B}(\tilde{B}'\tilde{B})^{-1}\tilde{B}'\tilde{q}}{w'w} \\
= & -\frac{z'M\tilde{q}}{w'w}(\tilde{B}'\tilde{B})^{-1}\tilde{B}'z \\
= & -\frac{\delta s}{w'w}(\tilde{B}'\tilde{B})^{-1}\tilde{B}'z,
\end{aligned} \tag{A.19}$$

and the last equality follows because $\delta^2 = \tilde{q}'PP'\tilde{q}$ and $s = \tilde{q}'PP'z/(\tilde{q}'PP'\tilde{q})^{\frac{1}{2}}$.

For the limiting distribution of $\hat{\gamma}_2$, we have

$$T^{-\frac{1}{2}}\hat{\gamma}_2 = \frac{(\sqrt{T}\hat{d}'\hat{U}^{-\frac{1}{2}})\hat{M}\hat{U}^{-\frac{1}{2}}q}{(\sqrt{T}\hat{d}'\hat{U}^{-\frac{1}{2}})\hat{M}(\sqrt{T}\hat{U}^{-\frac{1}{2}}\hat{d})} \xrightarrow{d} \frac{z'M\tilde{q}}{z'Mz} = \frac{\delta s}{w'w}. \tag{A.20}$$

This completes the proof of part (b) of Theorem 1.

Proof of Theorem 2.

part (a): Using Lemma A.1, we have

$$S = E[(\tilde{f}_t'\gamma_1^*)^2]U + BCB' \tag{A.21}$$

under the conditional homoskedasticity assumption. It follows that

$$\begin{aligned} V_u &= P'U^{-\frac{1}{2}}SU^{-\frac{1}{2}}P = E[(\tilde{f}_t'\gamma_1^*)^2]I_{N-K}, && \text{(A.22)} \\ V_r &= (\tilde{B}'\tilde{B})^{-\frac{1}{2}}\tilde{B}'U^{-\frac{1}{2}}SU^{-\frac{1}{2}}\tilde{B}(\tilde{B}'\tilde{B})^{-\frac{1}{2}} = E[(\tilde{f}_t'\gamma_1^*)^2]I_K + (\tilde{B}'\tilde{B})^{\frac{1}{2}}C(\tilde{B}'\tilde{B})^{\frac{1}{2}}, && \text{(A.23)} \\ \text{Cov}[u, r'] &= P'U^{-\frac{1}{2}}SU^{-\frac{1}{2}}\tilde{B}(\tilde{B}'\tilde{B})^{-\frac{1}{2}} = 0_{(N-K)\times K}. && \text{(A.24)} \end{aligned}$$

Let $\tilde{u} = w'u/(w'V_u w)^{\frac{1}{2}} = E[(\tilde{f}_t'\gamma_1^*)^2]^{-\frac{1}{2}}w'u/(w'w)^{\frac{1}{2}}$. 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.14) to

$$\sqrt{T}(\hat{\gamma}_1 - \gamma_1^*) \xrightarrow{d} -E[(\tilde{f}_t'\gamma_1^*)^2]^{\frac{1}{2}}\frac{\tilde{u}}{(w'w)^{\frac{1}{2}}}(\tilde{B}'\tilde{B})^{-1}\tilde{B}'z + (\tilde{B}'\tilde{B})^{-\frac{1}{2}}r. \tag{A.25}$$

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

$$\hat{V}_m(\hat{\gamma}) = \frac{1}{T^2}\sum_{t=1}^{T}\hat{h}_t\hat{h}_t', \tag{A.26}$$

where

$$\hat{h}_t = (\hat{D}'\hat{U}^{-1}\hat{D})^{-1}\hat{D}'\hat{U}^{-1}\hat{e}_t + (\hat{D}'\hat{U}^{-1}\hat{D})^{-1}([\tilde{f}_t',\ g_t]' - \hat{D}'\hat{U}^{-1}x_t)\hat{u}_t, \tag{A.27}$$

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

$$\begin{aligned} \hat{D}'\hat{U}^{-1} &= \begin{bmatrix} \tilde{B}'U^{-\frac{1}{2}} + O_p(T^{-\frac{1}{2}}) \\ \frac{1}{\sqrt{T}}z'U^{-\frac{1}{2}} + O_p(T^{-1}) \end{bmatrix}, && \text{(A.28)} \\ \hat{D}'\hat{U}^{-1}\hat{D} &= \begin{bmatrix} \tilde{B}'\tilde{B} + O_p(T^{-\frac{1}{2}}) & \frac{1}{\sqrt{T}}\tilde{B}'z + O_p(T^{-1}) \\ \frac{1}{\sqrt{T}}z'\tilde{B} + O_p(T^{-1}) & \frac{z'z}{T} + O_p(T^{-\frac{3}{2}}) \end{bmatrix}. && \text{(A.29)} \end{aligned}$$

Then, using the partitioned matrix inverse formula, we have

$$(\hat{D}'\hat{U}^{-1}\hat{D})^{-1} = \begin{bmatrix} H + O_p(T^{-\frac{1}{2}}) & -\sqrt{T}\frac{(\tilde{B}'\tilde{B})^{-1}\tilde{B}'z}{w'w} + O_p(1) \\ -\sqrt{T}\frac{z'\tilde{B}(\tilde{B}'\tilde{B})^{-1}}{w'w} + O_p(1) & \frac{T}{w'w} + O_p(T^{\frac{1}{2}}) \end{bmatrix}, \tag{A.30}$$

where

$$H = (\tilde{B}'[I_N - z(z'z)^{-1}z']\tilde{B})^{-1} = (\tilde{B}'\tilde{B})^{-1} + \frac{(\tilde{B}'\tilde{B})^{-1}\tilde{B}'zz'\tilde{B}(\tilde{B}'\tilde{B})^{-1}}{w'w}. \tag{A.31}$$

After simplification, we obtain

$$(\hat{D}'\hat{U}^{-1}\hat{D})^{-1}\hat{D}'\hat{U}^{-1} = \begin{bmatrix} (\tilde{B}'\tilde{B})^{-1}\tilde{B}'U^{-\frac{1}{2}} - \frac{(\tilde{B}'\tilde{B})^{-1}\tilde{B}'zw'P'U^{-\frac{1}{2}}}{w'w} + O_p(T^{-\frac{1}{2}}) \\ \frac{\sqrt{T}w'P'U^{-\frac{1}{2}}}{w'w} + O_p(1) \end{bmatrix}. \tag{A.32}$$

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

$$\hat{e} = \hat{D}\hat{\gamma} - q = \hat{D}(\hat{D}'\hat{U}^{-1}\hat{D})^{-1}\hat{D}'\hat{U}^{-1}q - q. \tag{A.33}$$

Using (A.15), (A.17), and the identity

$$I_N - \hat{U}^{-\frac{1}{2}}\hat{D}(\hat{D}'\hat{U}^{-1}\hat{D})^{-1}\hat{D}'\hat{U}^{-\frac{1}{2}} = \hat{M} - \hat{M}\hat{U}^{-\frac{1}{2}}\hat{d}(\hat{d}'\hat{U}^{-\frac{1}{2}}\hat{M}\hat{U}^{-\frac{1}{2}}\hat{d})^{-1}\hat{d}'\hat{U}^{-\frac{1}{2}}\hat{M}, \tag{A.34}$$

we can obtain the limiting distribution of $-\sqrt{T}\hat{U}^{-\frac{1}{2}}\hat{e}$ as

$$-\sqrt{T}\hat{U}^{-\frac{1}{2}}\hat{e} = \sqrt{T}\hat{M}\hat{U}^{-\frac{1}{2}}q - \sqrt{T}\hat{M}\hat{U}^{-\frac{1}{2}}\hat{d}\hat{\gamma}_2 \xrightarrow{d} My - Mz\frac{w'u}{w'w} = P\left(I_{N-K} - \frac{ww'}{w'w}\right)u, \tag{A.35}$$

and we have

$$\sqrt{T}\hat{u}_t \xrightarrow{d} -u'\left(I_{N-K} - \frac{ww'}{w'w}\right)P'U^{-\frac{1}{2}}x_t. \tag{A.36}$$

Using (A.30), (A.32), (A.36), and the fact that

$$\hat{e}_t = x_t(\tilde{f}_t'\hat{\gamma}_1 + \hat{\gamma}_2 g_t) - q = x_t\tilde{f}_t'\gamma_1^* - q + \frac{w'u}{w'w}x_t g_t + O_p(T^{-\frac{1}{2}}) \tag{A.37}$$

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

$$\begin{aligned} \hat{h}_{1t} \quad \xrightarrow{d} \quad & \left[(\tilde{B}'\tilde{B})^{-1}\tilde{B}'U^{-\frac{1}{2}} - \frac{(\tilde{B}'\tilde{B})^{-1}\tilde{B}'zw'P'U^{-\frac{1}{2}}}{w'w}\right]\left(x_t\tilde{f}_t'\gamma_1^* - q + x_t g_t\frac{w'u}{w'w}\right) \\ & + \frac{(\tilde{B}'\tilde{B})^{-1}\tilde{B}'z}{w'w}u'\left(I_{N-K} - \frac{ww'}{w'w}\right)P'U^{-\frac{1}{2}}x_t g_t, \end{aligned} \tag{A.38}$$

$$\frac{\hat{h}_{2t}}{\sqrt{T}} \xrightarrow{d} \frac{1}{w'w}w'P'U^{-\frac{1}{2}}\left(x_t\tilde{f}_t'\gamma_1^* - q + x_t g_t\frac{w'u}{w'w}\right) - \frac{1}{w'w}u'\left(I_{N-K} - \frac{ww'}{w'w}\right)P'U^{-\frac{1}{2}}x_t g_t. \tag{A.39}$$

Under the conditional homoskedasticity assumption, we have

$$\frac{1}{T}\sum_{t=1}^{T}(x_t\tilde{f}_t'\gamma_1^* - q)(x_t\tilde{f}_t'\gamma_1^* - q)' \xrightarrow{\text{a.s.}} S = E[(\tilde{f}_t'\gamma_1^*)^2]U + BCB'. \tag{A.40}$$

Together with the fact that

$$\frac{1}{T}\sum_{t=1}^{T} x_t x_t' g_t^2 \xrightarrow{\text{a.s.}} E[x_t x_t' g_t^2] = E[x_t x_t'] E[g_t^2] = U, \tag{A.41}$$

we can show that the estimated misspecification-robust covariance matrix of $\hat{\gamma}_1$ has a limiting distribution of

$$\begin{aligned} T\hat{V}_m(\hat{\gamma}_1) &= \frac{1}{T}\sum_{t=1}^{T} \hat{h}_{1t}\hat{h}_{1t}' \\ &\xrightarrow{d} E[(\tilde{f}_t'\gamma_1^*)^2]\left(1+\frac{\tilde{u}^2}{w'w}\right)\left[(\tilde{B}'\tilde{B})^{-1} + \frac{(\tilde{B}'\tilde{B})^{-1}\tilde{B}'zz'\tilde{B}(\tilde{B}'\tilde{B})^{-1}}{w'w}\right] + C \\ &\quad + u'\left(I_{N-K} - \frac{ww'}{w'w}\right) u \frac{(\tilde{B}'\tilde{B})^{-1}\tilde{B}'zz'\tilde{B}(\tilde{B}'\tilde{B})^{-1}}{(w'w)^2}. \end{aligned} \tag{A.42}$$

Let b_i be the i-th diagonal element of $(\tilde{B}'\tilde{B})^{-1}$. Then, we can readily show that

$$\tilde{z}_i = -\frac{\boldsymbol{\iota}_i'(\tilde{B}'\tilde{B})^{-1}\tilde{B}'z}{\sqrt{b_i}} \sim N(0,1), \tag{A.43}$$

$$v = \frac{u'[I_{N-K} - w(w'w)^{-1}w']u}{E[\tilde{f}_t\gamma_1^*)^2]} \sim \chi^2_{N-K-1}, \tag{A.44}$$

and v is independent of $\tilde{u}$, z and w. Using $\tilde{z}_i$ and v, we can express the limiting distribution of $s_m^2(\hat{\gamma}_{1i})$ as

$$Ts_m^2(\hat{\gamma}_{1i}) = T\boldsymbol{\iota}_i'\hat{V}_m(\hat{\gamma}_1)\boldsymbol{\iota}_i \xrightarrow{d} E[(\tilde{f}_t'\gamma_1^*)^2]b_i\left[\left(1+\frac{\tilde{u}^2}{w'w}\right)\left(1+\frac{\tilde{z}_i^2}{w'w}\right) + \frac{\tilde{z}_i^2 v}{(w'w)^2}\right] + c_i, \tag{A.45}$$

where c_i is the i-th diagonal element of C. In addition, by letting

$$\tilde{r}_i = (E[(\tilde{f}_t'\gamma_1^*)^2]b_i + c_i)^{-\frac{1}{2}}\boldsymbol{\iota}_i'(\tilde{B}'\tilde{B})^{-\frac{1}{2}}r \sim N(0,1), \tag{A.46}$$

we can write the i-th element in (A.25) as

$$\sqrt{T}(\hat{\gamma}_{1i} - \gamma_{1i}^*) \xrightarrow{d} (E[(\tilde{f}_t'\gamma_1^*)^2]b_i)^{\frac{1}{2}}\frac{\tilde{u}\tilde{z}_i}{(w'w)^{\frac{1}{2}}} + (E[(\tilde{f}_t'\gamma_1^*)^2]b_i + c_i)^{\frac{1}{2}}\tilde{r}_i. \tag{A.47}$$

Finally, by letting[14]

$$\lambda_i = 1 + \frac{c_i}{E[(\tilde{f}_t'\gamma_1^*)^2]b_i} > 0, \tag{A.48}$$

we can write the limiting distribution of $t_m(\hat{\gamma}_{1i})$ as

$$t_m(\hat{\gamma}_{1i}) = \frac{\hat{\gamma}_{1i} - \gamma_{1i}^*}{s_m(\hat{\gamma}_{1i})} \xrightarrow{d} \frac{\tilde{u}\tilde{z}_i + \sqrt{\lambda_i}\sqrt{w'w}\tilde{r}_i}{\left[\lambda_i(w'w) + \tilde{z}_i^2 + \tilde{u}^2\left(1+\frac{\tilde{z}_i^2}{w'w}\right) + \frac{\tilde{z}_i^2 v}{w'w}\right]^{\frac{1}{2}}}. \tag{A.49}$$

[14]From (A.46), we can see that $E[(\tilde{f}_t'\gamma_1^*)^2]b_i + c_i$ is the variance of $\boldsymbol{\iota}_i'(\tilde{B}'\tilde{B})^{-\frac{1}{2}}r$. Therefore, we have $\lambda_i > 0$.

The estimated covariance matrix of $\hat{\gamma}_1$ that assumes a correctly specified model is obtained by dropping the second term in (A.42). Then, it can be shown that

$$Ts_c^2(\hat{\gamma}_{1i}) \xrightarrow{d} E[(\tilde{f}_t'\gamma_1^*)^2]b_i\left[\left(1+\frac{\tilde{u}^2}{w'w}\right)\left(1+\frac{\tilde{z}_i^2}{w'w}\right)\right]+c_i \tag{A.50}$$

and hence

$$t_c(\hat{\gamma}_{1i}) = \frac{\hat{\gamma}_{1i}-\gamma_{1i}^*}{s_c(\hat{\gamma}_{1i})} \xrightarrow{d} \frac{\tilde{u}\tilde{z}_i+\sqrt{\lambda_i}\sqrt{w'w}\tilde{r}_i}{\left[\lambda_i(w'w)+\tilde{z}_i^2+\tilde{u}^2\left(1+\frac{\tilde{z}_i^2}{w'w}\right)\right]^{\frac{1}{2}}}. \tag{A.51}$$

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

$$\hat{\gamma}_2 \xrightarrow{d} \frac{w'u}{w'w} = \frac{(w'V_uw)^{\frac{1}{2}}}{(w'w)}\tilde{u}, \tag{A.52}$$

where $\tilde{u} = w'u/(w'V_uw)^{\frac{1}{2}} \sim N(0,1)$, and it is independent of w. Using (A.39), we obtain

$$\begin{aligned} s_m^2(\hat{\gamma}_2) &= \frac{1}{T^2}\sum_{t=1}^{T}\hat{h}_{2t}^2 \\ &\xrightarrow{d} \frac{1}{(w'w)^2}\left[w'V_uw+\frac{(w'u)^2}{w'w}\right]+\frac{u'[I_{N-K}-w(w'w)^{-1}w']u}{(w'w)^2} \\ &= \frac{w'V_uw+u'u}{(w'w)^2}. \end{aligned} \tag{A.53}$$

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

$$t_m(\hat{\gamma}_2) = \frac{\hat{\gamma}_2}{s_m(\hat{\gamma}_2)} \xrightarrow{d} \frac{\tilde{u}}{\left(1+\frac{u'u}{w'V_uw}\right)^{\frac{1}{2}}}. \tag{A.54}$$

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

$$s_c^2(\hat{\gamma}_2) \xrightarrow{d} \frac{1}{(w'w)^2}\left[w'V_uw+\frac{(w'u)^2}{w'w}\right] = \frac{w'V_uw}{(w'w)^2}\left(1+\frac{\tilde{u}^2}{w'w}\right). \tag{A.55}$$

It follows that

$$t_c(\hat{\gamma}_2) = \frac{\hat{\gamma}_2}{s_c(\hat{\gamma}_2)} \xrightarrow{d} \frac{\tilde{u}}{\left(1+\frac{\tilde{u}^2}{w'w}\right)^{\frac{1}{2}}}. \tag{A.56}$$

Under the conditional homoskedasticity assumption, $V_u = E[(\tilde{f}_t'\gamma_1^*)^2]I_K$, so we can write

$$t_m(\hat{\gamma}_2) \xrightarrow{d} \frac{\tilde{u}}{\left(1+\frac{\tilde{u}^2+v}{w'w}\right)^{\frac{1}{2}}}, \tag{A.57}$$

where v is defined in (A.44). This completes the proof of part (a) of Theorem 2.

part (b): We first derive the limiting distribution of $\hat{h}_t$ in (A.27). When a model is misspecified, we can see from part (b) of Theorem 1 that $\hat{\gamma}_2 = O_p(T^{\frac{1}{2}})$ and $\hat{\gamma}_1 = O_p(1)$, so $\hat{\gamma}_2$ is the dominant term. Therefore, using (27), we have

$$\hat{e}_t = x_t(\tilde{f}_t'\hat{\gamma}_1 + g_t\hat{\gamma}_2) - q = x_t g_t\hat{\gamma}_2 + O_p(1) = \frac{\sqrt{T}\delta s}{w'w}x_t g_t + O_p(1). \tag{A.58}$$

In addition, using (A.33), (A.34) and (A.20), we have

$$-\hat{U}^{-\frac{1}{2}}\hat{e} = \hat{M}\hat{U}^{-\frac{1}{2}}q - \hat{M}\hat{U}^{-\frac{1}{2}}\hat{d}\hat{\gamma}_2 \xrightarrow{d} M\tilde{q} - \frac{Mzz'M\tilde{q}}{z'Mz} = P[I_{N-K} - w(w'w)^{-1}w']P'\tilde{q}. \tag{A.59}$$

It follows that under a misspecified model,

$$\hat{u}_t = \hat{e}'\hat{U}^{-1}x_t \xrightarrow{d} -\tilde{q}'P[I_{N-K} - w(w'w)^{-1}w']P'U^{-\frac{1}{2}}x_t. \tag{A.60}$$

Then, using (A.30) and (A.32), we can express the limiting distribution of $\hat{h}_t = [\hat{h}_{1t}',\ \hat{h}_{2t}]'$ as

$$\begin{aligned}
\frac{\hat{h}_{1t}}{\sqrt{T}} &\xrightarrow{d} \frac{\tilde{q}'Pw}{w'w}(\tilde{B}'\tilde{B})^{-1}\tilde{B}'\left(I_N - \frac{zw'}{w'w}P'\right)U^{-\frac{1}{2}}x_t g_t \\
&\quad + \frac{(\tilde{B}'\tilde{B})^{-1}(\tilde{B}'z)}{w'w}\tilde{q}'P[I_{N-K} - w(w'w)^{-1}w']P'U^{-\frac{1}{2}}x_t g_t, \qquad \text{(A.61)} \\
\frac{\hat{h}_{2t}}{T} &\xrightarrow{d} \frac{\tilde{q}'Pw}{(w'w)^2}w'P'U^{-\frac{1}{2}}x_t g_t - \frac{1}{w'w}\tilde{q}'P[I_{N-K} - w(w'w)^{-1}w']P'U^{-\frac{1}{2}}x_t g_t. \qquad \text{(A.62)}
\end{aligned}$$

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

$$\tilde{B}'\left(I_N - \frac{zw'}{w'w}P'\right)P[I_{N-K} - w(w'w)^{-1}w']P'\tilde{q} = 0_K, \tag{A.63}$$

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

$$\begin{aligned}
\hat{V}_m(\hat{\gamma}_1) &= \frac{1}{T^2}\sum_{t=1}^{T}\hat{h}_{1t}\hat{h}_{1t}' \\
&= \frac{(\tilde{q}'Pw)^2}{(w'w)^2}\left[(\tilde{B}'\tilde{B})^{-1} + \frac{(\tilde{B}'\tilde{B})^{-1}\tilde{B}'zz'\tilde{B}(\tilde{B}'\tilde{B})^{-1}}{w'w}\right] \\
&\quad + \frac{1}{(w'w)^2}\left[\tilde{q}'PP'\tilde{q} - \frac{(\tilde{q}'Pw)^2}{w'w}\right](\tilde{B}'\tilde{B})^{-1}\tilde{B}'zz'\tilde{B}(\tilde{B}'\tilde{B})^{-1} \\
&= \frac{\delta^2}{(w'w)^2}\left[s^2(\tilde{B}'\tilde{B})^{-1} + (\tilde{B}'\tilde{B})^{-1}\tilde{B}'zz'\tilde{B}(\tilde{B}'\tilde{B})^{-1}\right]. \qquad \text{(A.64)}
\end{aligned}$$

Using $\tilde{z}_i$ as defined in (A.43), we can express the limiting distribution of $s_m^2(\hat{\gamma}_{1i})$ as

$$s_m^2(\hat{\gamma}_{1i}) = \boldsymbol{\iota}_i' \hat{V}_m(\hat{\gamma}_1)\boldsymbol{\iota}_i \xrightarrow{d} \frac{\delta^2 b_i}{(w'w)^2}(s^2 + \tilde{z}_i^2). \tag{A.65}$$

In addition, we can also use $\tilde{z}_i$ to express the i-th element in (26) as:

$$\hat{\gamma}_{1i} - \gamma_{1i}^* \xrightarrow{d} \frac{\delta s\sqrt{b_i}\tilde{z}_i}{w'w}. \tag{A.66}$$

It follows that when the model is misspecified, $t_m(\hat{\gamma}_{1i})$ has the following limiting distribution:

$$t_m(\hat{\gamma}_{1i}) = \frac{\hat{\gamma}_{1i} - \gamma_{1i}^*}{s_m(\hat{\gamma}_{1i})} \xrightarrow{d} \frac{s\tilde{z}_i}{\sqrt{s^2 + \tilde{z}_i^2}}. \tag{A.67}$$

To show that $t_m(\hat{\gamma}_{1i}) \xrightarrow{d} N(0, 1/4)$, 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 (ω, θ) is given by

$$f(\omega, \theta) = \frac{\omega e^{-\frac{\omega^2}{2}}}{2\pi} I_{\{\omega>0\}} I_{\{0<\theta<2\pi\}}. \tag{A.68}$$

Therefore, ω and θ are independent. Using the polar transformation, we obtain

$$\frac{s\tilde{z}_i}{\sqrt{s^2 + \tilde{z}_i^2}} = \omega\cos(\theta)\sin(\theta) = \frac{\omega\sin(2\theta)}{2}. \tag{A.69}$$

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

$$t_m(\hat{\gamma}_{1i}) \xrightarrow{d} N\left(0, \frac{1}{4}\right). \tag{A.70}$$

The estimated covariance matrix of $\hat{\gamma}_1$ that assumes a correctly specified model is obtained by dropping the second term in the line before (A.64). We can then show that

$$s_c^2(\hat{\gamma}_{1i}) \xrightarrow{d} \frac{\delta^2 s^2 b_i}{(w'w)^2}\left(1 + \frac{\tilde{z}_i^2}{w'w}\right). \tag{A.71}$$

Using (A.66), we can then obtain the limiting distribution of $t_c(\hat{\gamma}_{1i})$ as

$$t_c(\hat{\gamma}_{1i}) = \frac{\hat{\gamma}_{1i} - \gamma_{1i}^*}{s_c(\hat{\gamma}_{1i})} \xrightarrow{d} \frac{\tilde{z}_i}{\left(1 + \frac{\tilde{z}_i^2}{w'w}\right)^{\frac{1}{2}}}. \tag{A.72}$$

Turing our attention to the limiting distributions of $t_c(\hat{\gamma}_2)$ and $t_m(\hat{\gamma}_2)$, we use (A.62) and the fact that $\delta^2 = \tilde{q}'PP'\tilde{q}$ to obtain

$$\begin{aligned}\frac{s_m^2(\hat{\gamma}_2)}{T} &= \frac{1}{T^3}\sum_{t=1}^{T}\hat{h}_{2t}^2 \\ &\xrightarrow{d} \frac{(\tilde{q}'Pw)^2}{(w'w)^4}w'w + \frac{1}{(w'w)^2}\tilde{q}'P\left(I_{N-K} - \frac{ww'}{w'w}\right)P'\tilde{q} \\ &= \frac{\delta^2}{(w'w)^2}. \end{aligned} \tag{A.73}$$

Therefore, using (27), the t-statistic of $\hat{\gamma}_2$ under the misspecification-robust standard error is given by

$$t_m(\hat{\gamma}_2) = \frac{\hat{\gamma}_2}{s_m(\hat{\gamma}_2)} \xrightarrow{d} s \sim N(0,1). \tag{A.74}$$

For $s_c^2(\hat{\gamma}_2)$ which assumes a correctly specified model, we drop the second term of $\hat{h}_{2t}$ in (A.62), and we obtain

$$\frac{s_c^2(\hat{\gamma}_2)}{T} \xrightarrow{d} \frac{(\tilde{q}'Pw)^2}{(w'w)^3} = \frac{\delta^2 s^2}{(w'w)^3} \tag{A.75}$$

It follows that

$$t_c(\hat{\gamma}_2) = \frac{\hat{\gamma}_2}{s_c(\hat{\gamma}_2)} \xrightarrow{d} \text{sgn}(s)\sqrt{w'w}. \tag{A.76}$$

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 $w'w \sim \chi^2_{N-K}$.[15] 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 $t_c^2(\hat{\gamma}_{1i})$ and obvious for $t_m^2(\hat{\gamma}_{1i})$. First, comparing the limiting distribution of $t_c^2(\hat{\gamma}_{1i})$ with the limiting distribution of $t_m^2(\hat{\gamma}_{1i})$ 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 $t_m^2(\hat{\gamma}_{1i})$ is stochastically dominated by the limiting distribution of $t_c^2(\hat{\gamma}_{1i})$. It remains to be shown that the latter is stochastically dominated

[15] It is straightforward to show that the limiting probability density function of $t_c(\hat{\gamma}_2)$ is

$$f(t) = \frac{|t|^{N-K-1}e^{-\frac{t^2}{2}}}{2^{\frac{N-K}{2}}\Gamma\left(\frac{N-K}{2}\right)}. \tag{A.77}$$

by χ_1^2. From part (a) of Theorem 2, we have

$$t_c^2(\hat{\gamma}_{1i}) \xrightarrow{d} \frac{(\tilde{u}\tilde{z}_i + \sqrt{\lambda_i}\sqrt{w'w}\tilde{r}_i)^2}{\lambda_i(w'w) + \tilde{z}_i^2 + \tilde{u}^2\left(1 + \frac{\tilde{z}_i^2}{w'w}\right)}. \tag{A.78}$$

Let $\tilde{t} = \tilde{z}_i/\sqrt{w'w}$. It is easy to see that the limit of $t_c^2(\hat{\gamma}_{1i})$ is stochastically dominated by $(\tilde{t}\tilde{u} + \sqrt{\lambda_i}\tilde{r}_i)^2/(\lambda_i + \tilde{t}^2) \sim \chi_1^2$.

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 $t_c^2(\hat{\gamma}_2)$ and $t_m^2(\hat{\gamma}_2)$ are stochastically dominated by $\tilde{u}^2 \sim \chi_1^2$. This completes the proof of Corollary 1.

Proof of Theorem 3.

part (a): Using (A.35) in the proof of Theorem 2, we can easily obtain

$$T\hat{\delta}^2 = T\hat{e}'\hat{U}^{-1}\hat{e} \xrightarrow{d} u'[I_{N-K} - w(w'w)^{-1}w']u = u'P_wP_w'u, \tag{A.79}$$

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 $\tilde{v} = (P_w'V_uP_w)^{-\frac{1}{2}}P_w'u \sim N(0_{N-K-1}, I_{N-K-1})$, which is independent of w. Then, we have

$$T\hat{\delta}^2 \xrightarrow{d} \tilde{v}'(P_w'V_uP_w)\tilde{v}. \tag{A.80}$$

For testing $H_0 : \delta = 0$, $T\hat{\delta}^2$ is compared with $\sum_{i=1}^{N-K-1}\hat{\xi}_iX_i$, 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

$$\hat{S}^{\frac{1}{2}}\hat{U}^{-1}\hat{S}^{\frac{1}{2}} - \hat{S}^{\frac{1}{2}}\hat{U}^{-1}\hat{D}(\hat{D}'\hat{U}^{-1}\hat{D})^{-1}\hat{D}'\hat{U}^{-1}\hat{S}^{\frac{1}{2}}. \tag{A.81}$$

Using (A.34), we can write the above matrix as

$$\begin{aligned} & \hat{S}^{\frac{1}{2}}\hat{U}^{-\frac{1}{2}}[I_N - \hat{U}^{-\frac{1}{2}}\hat{D}(\hat{D}'\hat{U}^{-1}\hat{D})^{-1}\hat{D}'\hat{U}^{-\frac{1}{2}}]\hat{U}^{-\frac{1}{2}}\hat{S}^{\frac{1}{2}} \\ = \; & \hat{S}^{\frac{1}{2}}\hat{U}^{-\frac{1}{2}}\hat{M}\hat{U}^{-\frac{1}{2}}\hat{S}^{\frac{1}{2}} - \hat{S}^{\frac{1}{2}}\hat{U}^{-\frac{1}{2}}\hat{M}\hat{U}^{-\frac{1}{2}}\hat{d}(\hat{d}'\hat{U}^{-\frac{1}{2}}\hat{M}\hat{U}^{-\frac{1}{2}}\hat{d})^{-1}\hat{d}'\hat{U}^{-\frac{1}{2}}\hat{M}\hat{U}^{-\frac{1}{2}}\hat{S}^{\frac{1}{2}}. \end{aligned} \tag{A.82}$$

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 $\hat{P}_w\hat{P}_w' = I_{N-K} - \hat{P}'U^{-\frac{1}{2}}\hat{d}(\hat{d}'\hat{U}^{-\frac{1}{2}}\hat{M}\hat{U}^{-\frac{1}{2}}\hat{d})^{-1}\hat{d}'\hat{U}^{-\frac{1}{2}}\hat{P}$. We can easily show that $\hat{\xi}_i$'s are the nonzero eigenvalues of

$$\hat{S}^{\frac{1}{2}}\hat{U}^{-\frac{1}{2}}\hat{P}\hat{P}_w\hat{P}_w'\hat{P}'\hat{U}^{-\frac{1}{2}}\hat{S}^{\frac{1}{2}}, \tag{A.83}$$

or equivalently the eigenvalues of

$$\hat{P}_w'\hat{P}'\hat{U}^{-\frac{1}{2}}\hat{S}\hat{U}^{-\frac{1}{2}}\hat{P}\hat{P}_w. \tag{A.84}$$

Using (A.37), we can show that

$$\hat{P}'\hat{U}^{-\frac{1}{2}}\hat{e}_t \xrightarrow{d} P'U^{-\frac{1}{2}}e_t(\gamma^*) + \frac{w'u}{w'w}P'U^{-\frac{1}{2}}x_t g_t. \tag{A.85}$$

It follows that

$$\hat{P}'\hat{U}^{-\frac{1}{2}}\hat{S}\hat{U}^{-\frac{1}{2}}\hat{P} \xrightarrow{d} P'U^{-\frac{1}{2}}SU^{-\frac{1}{2}}P + \frac{(w'u)^2}{(w'w)^2}I_{N-K} = V_u + \frac{(w'V_u w)\tilde{u}^2}{(w'w)^2}I_{N-K}, \tag{A.86}$$

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

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

$$T\hat{\delta}^2 \xrightarrow{d} E[(\tilde{f}_t'\gamma_1^*)^2]\tilde{v}'\tilde{v} \sim E[(\tilde{f}_t'\gamma_1^*)^2]\chi^2_{N-K-1}, \tag{A.87}$$

$$\hat{P}_w'\hat{P}'\hat{U}^{-\frac{1}{2}}\hat{S}\hat{U}^{-\frac{1}{2}}\hat{P}\hat{P}_w \xrightarrow{d} E[(\tilde{f}_t'\gamma_1^*)^2]\left(1 + \frac{\tilde{u}^2}{w'w}\right)I_{N-K-1}. \tag{A.88}$$

It follows that

$$\hat{\xi}_i \xrightarrow{d} E[(\tilde{f}_t'\gamma_1^*)^2]\left(1 + \frac{\tilde{u}^2}{w'w}\right) = \frac{E[(\tilde{f}_t'\gamma_1^*)^2]}{Q_1}, \tag{A.89}$$

where $Q_1 = w'w/(\tilde{u}^2 + w'w) \sim \text{Beta}\left(\frac{N-K}{2}, \frac{1}{2}\right)$ and it is independent of $\tilde{v}'\tilde{v}$. Therefore, the limiting probability of rejection of the HJ-distance test of size α is

$$\int_0^1 P\left[\chi^2_{N-K-1} > \frac{c_\alpha}{q}\right] f_{Q_1}(q)\mathrm{d}q, \tag{A.90}$$

where c_α is the $100(1-\alpha)$ percentile of χ^2_{N-K-1}. Since $0 < Q_1 < 1$, the limiting probability of rejection is less than α. This completes the proof of part (a) of Theorem 3.

part (b): Using (A.59), the limiting distribution of the squared sample HJ-distance $\hat{\delta}^2 = \hat{e}'\hat{U}^{-1}\hat{e}$ can be obtained as

$$\begin{aligned}\hat{\delta}^2 &\xrightarrow{d} \tilde{q}'P[I_{N-K} - w(w'w)^{-1}w']P\tilde{q} \\ &= (\tilde{q}'PP'\tilde{q})\frac{w'[I_{N-K} - P'\tilde{q}(\tilde{q}'PP'\tilde{q})^{-1}\tilde{q}'P]w}{w'w} = \delta^2 Q_2,\end{aligned} \tag{A.91}$$

where

$$Q_2 = \frac{w'[I_{N-K} - P'\tilde{q}(\tilde{q}'PP'\tilde{q})^{-1}\tilde{q}'P]w}{w'w} \sim \text{Beta}\left(\frac{N-K-1}{2}, \frac{1}{2}\right) \tag{A.92}$$

and it is independent of w.

From the proof of part (a), we know $\hat{\xi}_i$'s are the eigenvalues of

$$\hat{P}_w'\hat{P}'\hat{U}^{-\frac{1}{2}}\hat{S}\hat{U}^{-\frac{1}{2}}\hat{P}\hat{P}_w. \tag{A.93}$$

From (35) and (27), we have

$$\frac{\hat{S}}{T} \xrightarrow{d} \frac{\delta^2 s^2}{(w'w)^2}U, \tag{A.94}$$

which implies

$$\frac{\hat{P}_w'\hat{P}'\hat{U}^{-\frac{1}{2}}\hat{S}\hat{U}^{-\frac{1}{2}}\hat{P}\hat{P}_w}{T} \xrightarrow{d} \frac{\delta^2 s^2}{(w'w)^2}I_{N-K-1}. \tag{A.95}$$

Therefore,

$$\frac{\hat{\xi}_i}{T} \xrightarrow{d} \frac{\delta^2 s^2}{(w'w)^2} = \frac{\delta^2(1-Q_2)}{w'w}, \tag{A.96}$$

and 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

$$w'w > \frac{c_\alpha Q_2}{1-Q_2}. \tag{A.97}$$

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 α is

$$\int_0^1 P\left[\chi^2_{N-K} > \frac{c_\alpha q}{1-q}\right] f_{Q_2}(q)\mathrm{d}q. \tag{A.98}$$

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

References

Burnside, Craig A., 2010, Identification and inference in linear stochastic discount factor models, NBER Working paper 16634.

Burnside, Craig A., 2011, The cross-section of foreign currency risk premia and consumption growth risk: Comment, *American Economic Review* 101, 3456–3476.

Campbell, John Y., 1996, Understanding risk and return, *Journal of Political Economy* 104, 298–345.

Cochrane, John H., 1996, A cross-sectional test of an investment-based asset pricing model, *Journal of Political Economy* 104, 572–621.

Cragg, John G., and Stephen G. Donald, 1997, Inferring the rank of a matrix, *Journal of Econometrics* 76, 223–250.

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

Fama, Eugene F., and James D. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 71, 607–636.

Gospodinov, Nikolay, Raymond Kan, and Cesare Robotti, 2012a, Further results on the limiting distribution of GMM sample moment conditions, *Journal of Business and Economic Statistics*, forthcoming.

Gospodinov, Nikolay, Raymond Kan, and Cesare Robotti, 2012b, Chi-squared tests for evaluation and comparison of asset pricing models, Working paper, Concordia University.

Hall, Alastair R., and Atsushi Inoue, 2003, The large sample behaviour of the generalized method of moments estimator in misspecified models, *Journal of Econometrics* 114, 361–394.

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

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

Jagannathan, Ravi, and Zhenyu Wang, 1996, The conditional CAPM and the cross-section of expected returns, *Journal of Finance* 51, 3–53.

Jagannathan, Ravi, and Zhenyu Wang, 1998, An asymptotic theory for estimating beta-pricing models using cross-sectional regression, *Journal of Finance* 53, 1285–1309.

Jørgensen–Vissing, Annette, and Orazio P. Attanasio, 2003, Stock-market participation, intertemporal substitution, and risk-aversion, *American Economic Review Papers and Proceedings* 93, 383–391.

Kan, Raymond, and Cesare Robotti, 2008, Specification tests of asset pricing models using excess returns, *Journal of Empirical Finance* 15, 816–838.

Kan, Raymond, and Cesare Robotti, 2009, Model comparison using the Hansen-Jagannathan distance, *Review of Financial Studies* 22, 3449–3490.

Kan, Raymond, and Chu Zhang, 1999a, Two-pass tests of asset pricing models with useless factors, *Journal of Finance* 54, 204–235.

Kan, Raymond, and Chu Zhang, 1999b, GMM tests of stochastic discount factor models with useless factors, *Journal of Financial Economics* 54, 103–127.

Khalaf, Lynda A., and Huntley Schaller, 2011, How Fama-MacBeth can go wrong – And an informative solution, Working paper, Carleton University.

Kleibergen, Frank, 2009, Tests of risk premia in linear factor models, *Journal of Econometrics* 149, 149–173.

Kleibergen, Frank, 2010, Reality checks for and of factor pricing, Working paper, Brown University.

Lettau, Martin, and Sydney C. Ludvigson, 2001, Resurrecting the (C)CAPM: A cross-sectional test when risk premia are time-varying, *Journal of Political Economy* 109, 1238–1287.

Li, Quing, Maria Vassalou, and Yuhang Xing, 2006, Sector investment growth rates and the cross-section of equity returns, *Journal of Business* 79, 1637–1665.

Lustig, Hanno, and Stijn Van Nieuwerburgh, 2005, Housing collateral, consumption insurance, and risk premia: An empirical perspective, *Journal of Finance* 60, 1167–1219.

Merton, Robert C., 1973, An intertemporal capital asset pricing model, *Econometrica* 41, 867–887.

Petkova, Ralitsa, 2006, Do the Fama-French factors proxy for innovations in predictive variables?, *Journal of Finance* 61, 581–612.

Santos, Tano, and Pietro Veronesi, 2006, Labor income and predictable stock returns, *Review of Financial Studies* 19, 1–44.

Yogo, Motohiro, 2006, A consumption-based explanation of expected stock returns, *Journal of Finance* 61, 539–580.

Table I
Empirical Size of the *t*-tests in a Model with One Useful Factor

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. γ_0 is the coefficient on the constant term and γ_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. The rejection rates for the limiting case ($T = \infty$) are based on the standard normal distribution.

Panel A: Correctly Specified Model

		$\hat{\gamma}_0$			$\hat{\gamma}_1$		
t-test	T	10%	5%	1%	10%	5%	1%
t_c	200	0.168	0.138	0.106	0.097	0.049	0.008
	600	0.134	0.095	0.059	0.100	0.049	0.009
	1,000	0.119	0.078	0.039	0.099	0.049	0.010
	∞	0.100	0.050	0.010	0.100	0.050	0.010
t_m	200	0.168	0.138	0.106	0.097	0.048	0.008
	600	0.134	0.095	0.059	0.100	0.049	0.009
	1,000	0.119	0.078	0.039	0.099	0.049	0.010
	∞	0.100	0.050	0.010	0.100	0.050	0.010

Panel B: Misspecified Model

		$\hat{\gamma}_0$			$\hat{\gamma}_1$		
t-test	T	10%	5%	1%	10%	5%	1%
t_c	200	0.167	0.137	0.105	0.099	0.048	0.009
	600	0.134	0.096	0.059	0.099	0.050	0.010
	1000	0.122	0.079	0.040	0.100	0.051	0.010
	∞	0.100	0.050	0.010	0.100	0.050	0.010
t_m	200	0.167	0.137	0.105	0.098	0.048	0.009
	600	0.134	0.096	0.059	0.099	0.050	0.010
	1000	0.122	0.078	0.040	0.100	0.051	0.010
	∞	0.100	0.050	0.010	0.100	0.050	0.010

Table II
Empirical Size of the *t*-tests in a Model with One Useless Factor

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. γ_0 is the coefficient on the constant term and γ_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 obtained based on the asymptotic distributions given in Theorem 2.

Panel A: Correctly Specified Model

		$\hat{\gamma}_0$			$\hat{\gamma}_1$		
t-test	T	10%	5%	1%	10%	5%	1%
t_c	200	0.003	0.001	0.000	0.103	0.045	0.005
	600	0.001	0.000	0.000	0.087	0.036	0.003
	1000	0.001	0.000	0.000	0.084	0.033	0.003
	∞	0.001	0.000	0.000	0.080	0.032	0.003
t_m	200	0.001	0.000	0.000	0.026	0.007	0.000
	600	0.000	0.000	0.000	0.019	0.005	0.000
	1000	0.000	0.000	0.000	0.017	0.004	0.000
	∞	0.000	0.000	0.000	0.016	0.004	0.000

Panel B: Misspecified Model

		$\hat{\gamma}_0$			$\hat{\gamma}_1$		
t-test	T	10%	5%	1%	10%	5%	1%
t_c	200	0.012	0.003	0.000	0.314	0.212	0.074
	600	0.017	0.004	0.000	0.484	0.389	0.215
	1000	0.022	0.005	0.000	0.569	0.485	0.317
	∞	0.080	0.032	0.003	1.000	1.000	1.000
t_m	200	0.001	0.000	0.000	0.067	0.028	0.003
	600	0.001	0.000	0.000	0.079	0.035	0.005
	1000	0.001	0.000	0.000	0.085	0.040	0.006
	∞	0.001	0.000	0.000	0.100	0.050	0.010

Table III
Empirical Size of the *t*-tests in a Model with a Useful and a Useless Factor

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. γ_0 is the coefficient on the constant term, γ_1 is the coefficient on the useful factor, and γ_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 obtained based on the asymptotic distributions given in Theorem 2.

Panel A: Correctly Specified Model

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

Panel B: Misspecified Model

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

Table IV
Empirical Size and Power of the HJ-Distance Test

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 II. The rejection rates for the limiting case ($T = \infty$) are obtained from the asymptotic results given in Theorem 3.

Panel A: Model with a Useful Factor

	$N = 11$						$N = 26$					
	Size			Power			Size			Power		
T	10%	5%	1%	10%	5%	1%	10%	5%	1%	10%	5%	1%
200	0.129	0.069	0.017	0.298	0.195	0.070	0.229	0.145	0.050	0.994	0.987	0.951
600	0.110	0.057	0.012	0.607	0.480	0.254	0.137	0.074	0.018	1.000	1.000	1.000
1000	0.108	0.056	0.012	0.829	0.736	0.509	0.121	0.064	0.014	1.000	1.000	1.000
∞	0.100	0.050	0.010	1.000	1.000	1.000	0.100	0.050	0.010	1.000	1.000	1.000

Panel B: Model with a Useless Factor

	$N = 11$						$N = 26$					
	Size			Power			Size			Power		
T	10%	5%	1%	10%	5%	1%	10%	5%	1%	10%	5%	1%
200	0.091	0.046	0.010	0.334	0.228	0.092	0.190	0.115	0.036	0.970	0.949	0.878
600	0.076	0.036	0.007	0.692	0.597	0.395	0.110	0.058	0.013	0.996	0.994	0.987
1000	0.073	0.035	0.006	0.830	0.774	0.639	0.097	0.050	0.010	0.998	0.997	0.993
∞	0.070	0.033	0.006	0.944	0.931	0.902	0.079	0.038	0.007	0.999	0.999	0.998

Panel C: Model with a Useful and a Useless Factor

	$N = 11$						$N = 26$					
	Size			Power			Size			Power		
T	10%	5%	1%	10%	5%	1%	10%	5%	1%	10%	5%	1%
200	0.090	0.047	0.010	0.200	0.122	0.039	0.184	0.111	0.035	0.959	0.932	0.845
600	0.076	0.037	0.007	0.423	0.311	0.142	0.109	0.057	0.013	0.995	0.992	0.984
1000	0.074	0.035	0.007	0.609	0.502	0.299	0.097	0.049	0.010	0.997	0.996	0.991
∞	0.069	0.033	0.006	0.926	0.910	0.875	0.079	0.038	0.007	0.999	0.998	0.997

Table V
Survival Rates of Risk Factors

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 γ_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.

Panel A: Correctly Specified Model

	Useful Factor ($\gamma_1^* \neq 0$)		Useful Factor ($\gamma_2^* = 0$)		Useless Factor	
T	$t_c(\hat{\gamma}_1)$	$t_m(\hat{\gamma}_1)$	$t_c(\hat{\gamma}_2)$	$t_m(\hat{\gamma}_2)$	$t_c(\hat{\gamma}_3)$	$t_m(\hat{\gamma}_3)$
200	0.377	0.365	0.038	0.035	0.036	0.005
600	0.853	0.842	0.040	0.037	0.029	0.004
1000	0.974	0.970	0.041	0.039	0.029	0.003

Panel B: Misspecified Model

	Useful Factor ($\gamma_1^* \neq 0$)		Useful Factor ($\gamma_2^* = 0$)		Useless Factor	
T	$t_c(\hat{\gamma}_1)$	$t_m(\hat{\gamma}_1)$	$t_c(\hat{\gamma}_2)$	$t_m(\hat{\gamma}_2)$	$t_c(\hat{\gamma}_3)$	$t_m(\hat{\gamma}_3)$
200	0.357	0.329	0.036	0.029	0.190	0.024
600	0.781	0.730	0.040	0.027	0.369	0.034
1000	0.907	0.862	0.040	0.022	0.470	0.039

Table VI
Monthly Analysis of Some Popular Linear Asset Pricing Models

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 ($\mathcal{W}^*$) 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 ($\mathcal{W}^*$) with the corresponding p-values (p-val) for each model. The t-tests under correct model specification and the model misspecification-robust t-tests 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.

Panel A: Rank Test for Individual Factors

Test	vw	smb	hml	c	c_{nd}	c_d	cay	$c{\cdot}cay$	$prem$	lab	$term$	def	div	rf
$\mathcal{W}^*$	165.1	167.2	155.6	39.4	33.6	37.3	37.3	22.6	32.1	33.6	30.9	38.6	144.7	21.8
p-val	0.000	0.000	0.000	0.034	0.118	0.054	0.054	0.599	0.156	0.116	0.193	0.040	0.000	0.649

Panel B: HJ-Distance, Lagrange Multiplier, and Rank Tests

Model	$\hat{\delta}$	p-val	LM	p-val	$\mathcal{W}^*$	p-val
CAPM	0.430	0.000	92.927	0.000	165.1	0.000
C-LAB	0.406	0.000	68.034	0.000	26.3	0.286
FF3	0.367	0.000	66.358	0.000	116.0	0.000
ICAPM	0.339	0.252	23.962	0.244	19.7	0.540
CCAPM	0.429	0.000	80.135	0.000	39.4	0.034
CC-CAY	0.395	0.001	50.374	0.001	23.1	0.454
D-CCAPM	0.408	0.000	70.547	0.000	30.4	0.139

Panel C: t-tests Using Standard Errors Under Correct Model Specification

Model	vw	smb	hml	c	c_{nd}	c_d	cay	$c{\cdot}cay$	$prem$	lab	$term$	def	div	rf
CAPM	−2.60													
C-LAB	−1.18								−1.33	2.51				
FF3	−3.67	−1.77	−5.31											
ICAPM	1.56										−2.55	2.05	1.67	0.58
CCAPM				−2.65										
CC-CAY				−2.22			−2.83	0.73						
D-CCAPM	−0.89				−2.18	−1.18								

Panel D: t-tests Using Model Misspecification-Robust Standard Errors

Model	vw	smb	hml	c	c_{nd}	c_d	cay	$c{\cdot}cay$	$prem$	lab	$term$	def	div	rf
CAPM	−2.60													
C-LAB	−1.03								−0.60	1.17				
FF3	−3.66	−1.77	−5.33											
ICAPM	1.16										−1.76	1.53	1.27	0.49
CCAPM				−1.58										
CC-CAY				−1.66			−1.84	0.45						
D-CCAPM	−0.72				−1.30	−0.79								

Table VII
Quarterly Analysis of Some Popular Linear Asset Pricing Models

The table presents the estimation results of seven 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 1959 Q2 until 2007 Q4. Panel A reports the rank restriction test ($\mathcal{W}^*$) 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 ($\mathcal{W}^*$) with the corresponding p-values (p-val) for each model. The t-tests under correct model specification and the model misspecification-robust t-tests 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.

Panel A: Rank Test for Individual Factors

Test	vw	$vw{\cdot}ml$	smb	hml	c	c_{nd}	c_d	cay	$c{\cdot}cay$	my	$c{\cdot}my$	i_h	i_c	i_{nc}
$\mathcal{W}^*$	70.6	70.6	93.0	67.0	32.2	29.2	38.5	35.0	20.7	32.4	20.3	32.8	25.7	32.6
p-val	0.000	0.000	0.000	0.000	0.153	0.257	0.041	0.089	0.710	0.147	0.730	0.137	0.422	0.142

Panel B: HJ-Distance, Lagrange Multiplier, and Rank Tests

Model	$\hat{\delta}$	p-val	LM	p-val	$\mathcal{W}^*$	p-val
CAPM	0.629	0.000	63.904	0.000	70.6	0.000
C-ML	0.608	0.000	63.995	0.000	26.7	0.318
FF3	0.549	0.000	49.269	0.001	66.4	0.000
CCAPM	0.658	0.000	66.084	0.000	32.2	0.153
CC-CAY	0.587	0.003	37.559	0.021	18.5	0.727
D-CCAPM	0.612	0.000	59.720	0.000	19.8	0.654
CC-MY	0.625	0.000	49.750	0.001	19.2	0.688
SIM	0.540	0.129	35.155	0.037	20.7	0.600

Panel C: t-tests Using Standard Errors Under Correct Model Specification

Model	vw	$vw{\cdot}ml$	smb	hml	c	c_{nd}	c_d	cay	$c{\cdot}cay$	my	$c{\cdot}my$	i_h	i_c	i_{nc}
CAPM	−2.94													
C-ML	2.07	−2.10												
FF3	−3.69		−0.28	−4.33										
CCAPM					−1.80									
CC-CAY					−0.62			−2.60	−0.38					
D-CCAPM	−1.79					−1.88	0.36							
CC-MY					−1.80					1.73	−0.01			
SIM												−2.90	2.50	−1.48

Panel D: t-tests Using Model Misspecification-Robust Standard Errors

Model	vw	$vw{\cdot}ml$	smb	hml	c	c_{nd}	c_d	cay	$c{\cdot}cay$	my	$c{\cdot}my$	i_h	i_c	i_{nc}
CAPM	−2.93													
C-ML	1.51	−1.54												
FF3	−3.71		−0.28	−4.37										
CCAPM					−1.27									
CC-CAY					−0.37			−1.60	−0.16					
D-CCAPM	−1.64					−1.17	0.25							
CC-MY					−1.31					1.06	−0.01			
SIM												−2.15	1.69	−1.10

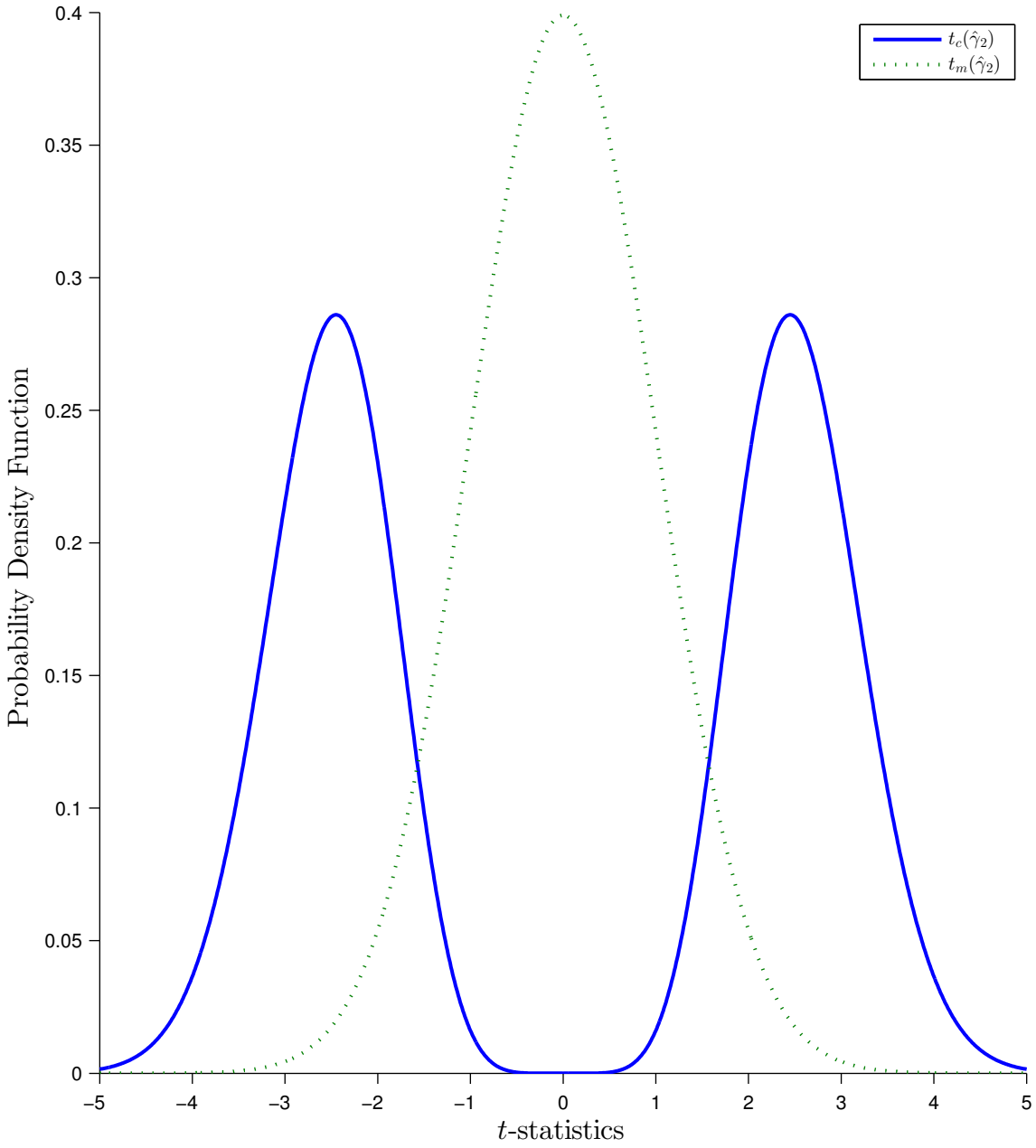

Figure 1. Asymptotic distributions of $t_c(\hat{\gamma}_2)$ and $t_m(\hat{\gamma}_2)$ under misspecified models. The figure presents the probability density functions of the limiting distributions of $t_c(\hat{\gamma}_2)$ and $t_m(\hat{\gamma}_2)$, the t-ratios for the useless factor that use standard errors constructed under correctly specified and potentially misspecified models, respectively, for $N - K = 7$ (see part (b) of Theorem 2).

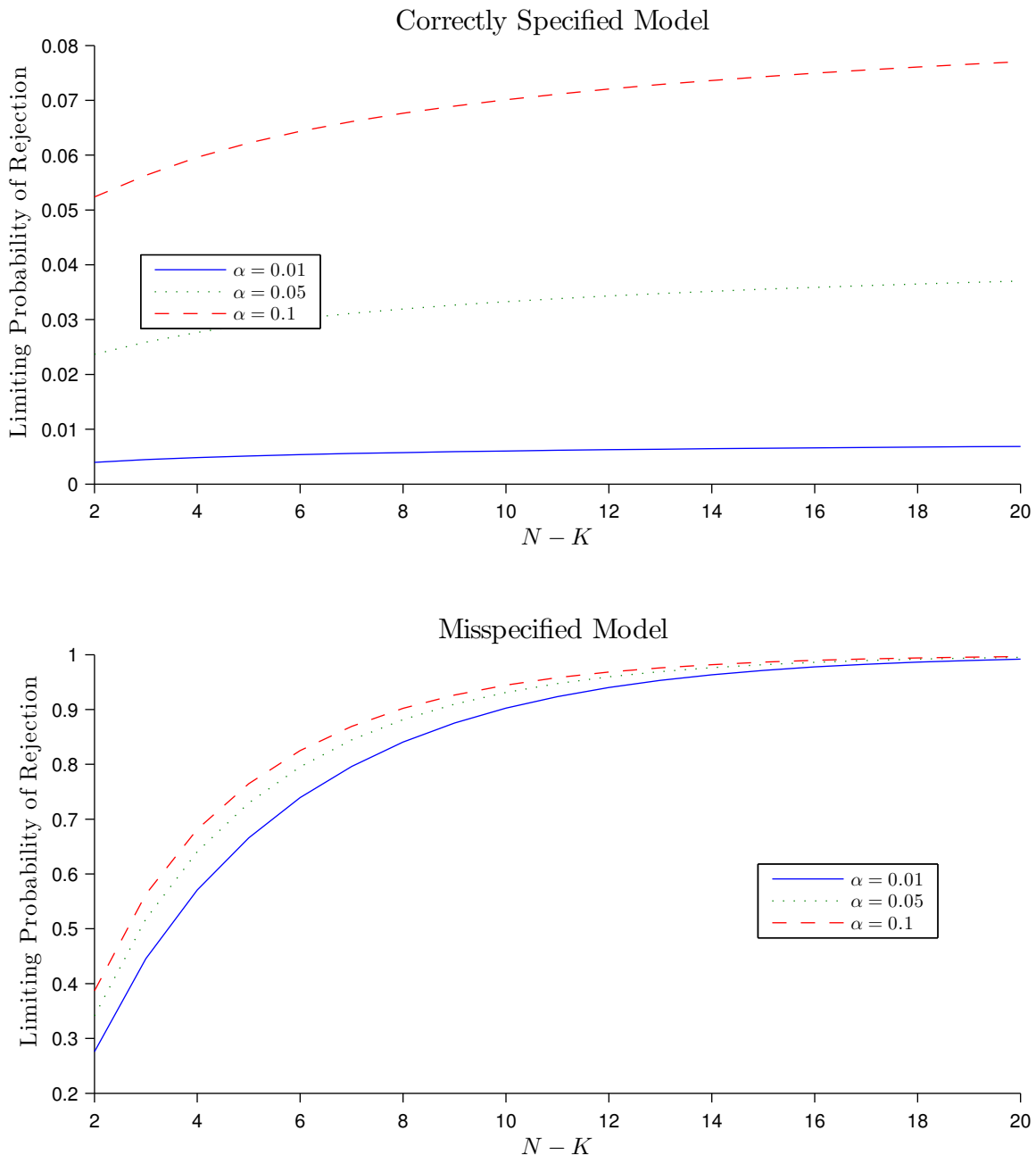

Figure 2. Limiting probabilities of rejection of the HJ-distance test. The figure presents the limiting probabilities of rejection of the HJ-distance test under correctly specified and misspecified models when one of the factors is useless.

Additional Empirical Results for

Robust Inference in Linear Asset Pricing Models

Nikolay Gospodinov, Raymond Kan, and Cesare Robotti

September 2012

In this appendix, we provide some additional empirical results that are not included in the paper.

A Carry Trade Portfolios and Consumption Growth

Lustig and Verdelhan (2007) claim that aggregate consumption growth risk explains the excess returns to borrowing U.S. dollars to finance lending in other currencies. In this section, we revisit their findings in light of our new theoretical and simulation results.

The return data are the annual excess returns on eight currency portfolios from 1953 until 2002.[1] It is well-known that when only excess returns are used in the analysis, the mean of the SDF cannot be identified. As a result, researchers have to choose some normalization of the SDF. We follow Kan and Robotti (2008) and employ the modified HJ-distance as our misspecification measure. Kan and Robotti (2008) show that the modified HJ-distance measure has the desirable property of being invariant to affine transformations of the factors. Under this metric and when the model is linear, the SDF is written as

$$y_t(\gamma) = 1 - \gamma'(f_t - E[f_t]).$$

It should be noted that the theory developed in the paper can be easily modified to accommodate the modified HJ-distance case. Importantly, under misspecified models, the misspecification-robust t-statistic will still have a standard normal limiting distribution.

We consider the same four linear asset pricing models analyzed in Table V of Lustig and Verdelhan (2007), and estimate them using the sample modified HJ-distance.[2] The first model considered is the consumption CAPM (CCAPM) with SDF given by

$$y_t^{CCAPM}(\gamma) = 1 - \gamma_1(c_{nd,t} - E[c_{nd,t}]),$$

where c_{nd} is the growth rate in real per capita nondurable consumption. The second model (DCAPM) explicitly breaks down consumption into consumption of nondurables and consumption of durables (c_d):

$$y_t^{DCAPM}(\gamma) = 1 - \gamma_1(c_{nd,t} - E[c_{nd,t}]) - \gamma_2(c_{d,t} - E[c_{d,t}]).$$

[1] See Lustig and Verdelhan (2007) for a detailed description of the data.

[2] Since Lustig and Verdelhan (2007) estimate the models using the ordinary least squares two-pass cross-sectional regression methodology, our results are not directly comparable with theirs.

The third model (EZ-CCAPM)

$$y_t^{EZ-CCAPM}(\gamma) = 1 - \gamma_1(c_{nd,t} - E[c_{nd,t}]) - \gamma_2(vw_t - E[vw_t])$$

is a linearized version of the model of Epstein and Zin (1989) with nondurable consumption and the market return (vw) as risk factors. The last model (D-CCAPM)

$$y_t^{D-CCAPM}(\gamma) = 1 - \gamma_1(c_{nd,t} - E[c_{nd,t}]) - \gamma_2(c_{d,t} - E[c_{d,t}]) - \gamma_3(vw_t - E[vw_t]),$$

is a linearized version of the durable consumption CAPM of Yogo (2006).

Table A.1 about here

Panel A of Table A.1 shows that for all risk factors, we cannot reject the null of $E[x_t(f_{it} - E[f_{it}])] = 0_N$. Similarly, for all four models, we cannot reject the null of deficient rank at any conventional significance level.[3] At the same time, none of the models is rejected by the specification test based on the modified HJ-distance at the 5% confidence level.[4] However, as our theoretical analysis suggests, the outcome of the specification test based on the HJ-distance should be interpreted with caution when useless factors are included in the model. In the useless factor case, the HJ-distance test will be inconsistent under the alternative and may have very low power in rejecting misspecified models.

In addition, Panel D clearly shows that no factor survives the model selection procedure based on misspecification-robust t-tests. Even the market factor does not appear to be priced, probably given its small (9%) average absolute correlation with the excess returns on the eight currency portfolios. Consistent with Burnside (2011), our analysis based on the modified HJ-distance shows that there is not enough statistical evidence for us to conclude that consumption growth risk prices the cross-section of carry trade portfolio returns.

B Momentum Portfolios and Industrial Production Growth

Liu and Zhang (2008) claim that the growth rate of industrial production explains the cross-sectional variation in momentum portfolios, yet the average absolute correlation between the mo-

[3] In this context, for a model to be identified, we need the covariance matrix between the excess returns on the test assets and the risk factors to be of full column rank.

[4] Using the LM test and a 5% significance level, we can reject the null of correct specification for CCAPM and EZ-CCAPM.

mentum portfolio returns and the growth rate of industrial production is only about 2%.

The return data are the monthly excess returns on ten momentum, ten size and ten book-to-market portfolios from January 1960 until December 2004.[5]

We consider the same four linear asset pricing models analyzed in Table V of Liu and Zhang (2008), and estimate them using the sample modified HJ-distance.[6] The first model considered is a one-factor model (MP1) with the growth rate of industrial production, mp, as the only factor. Its SDF is given by

$$y_t^{MP1}(\gamma) = 1 - \gamma_1(c_{mp,t} - E[c_{mp,t}]).$$

The second model (FF3) is the usual three-factor model of Fama and French (1993) with SDF given by

$$y_t^{FF3}(\gamma) = 1 - \gamma_1(vw_t - E[vw_t]) - \gamma_2(smb_t - E[smb_t]) - \gamma_3(hml_t - E[hml_t]),$$

where vw, smb and hml are the market, size, and book-to-market factors of Fama and French (1993), respectively. The third model (MP4) is an augmented FF3 model with mp as the additional factor. Its SDF is given by

$$y_t^{MP4}(\gamma) = 1 - \gamma_1(vw_t - E[vw_t]) - \gamma_2(smb_t - E[smb_t]) - \gamma_3(hml_t - E[hml_t]) - \gamma_4(mp_t - E[mp_t]).$$

The last model is the intertemporal CAPM-type model proposed by Chen, Roll and Ross (CRR, 1986):

$$\begin{aligned} y_t^{CRR}(\gamma) &= 1 - \gamma_1(ui_t - E[ui_t]) - \gamma_2(dei_t - E[dei_t]) - \gamma_3(uts_t - E[uts_t]) \\ &\quad -\gamma_4(upr_t - E[upr_t]) - \gamma_5(mp_t - E[mp_t]), \end{aligned}$$

where ui, dei, upr, and uts denote the unexpected inflation, the change in expected inflation, the term premium, and the default premium factors, respectively.

Table B.1 about here

Table B.1 shows that all four models do not pass the HJ-distance and LM tests. In addition, we cannot reject the null of deficient rank for MP4 and CRR at the 5% nominal level and only

[5] We refer to Liu and Zhang (2008) for a detailed description of the data.

[6] Since Liu and Zhang (2008) estimate the models using the ordinary least squares two-pass cross-sectional regression methodology, our results are not directly comparable with theirs.

FF3 appears to be identified at any conventional nominal level. The results of the specification and rank restriction tests clearly point to the need for robust statistical methods.

Consistent with the other empirical illustrations, all the non-traded factors do not survive the model selection procedure based on misspecification-robust t-tests. The only evidence of pricing comes from the market factor and the book-to-market factor of Fama and French (1993). In summary, there is not enough statistical evidence for us to conclude that industrial production risk prices the cross-section of momentum portfolio returns.

C The Stock-Watson Factors

The paper considers several models with macroeconomic variables as risk factors. It is possible that these individual factors do not capture adequately the risk incorporated in all of the macroeconomic data that is available to market participants. One approach to extract parsimoniously the common variation in macroeconomic variables is the factor analysis advanced by Stock and Watson (2002a, 2002b). See also Ludvigson and Ng (2007, 2009) for a similar approach in the analysis of stock and bond risk premia. In this section, we follow Stock and Watson (2005) and construct three orthogonal factors that summarize the dynamics of 127 macroeconomic time series for the period February 1959 – July 2007. We use the same variables and transformations as described in Stock and Watson (2005) with the exception of four series (houses authorized by building permits for Northeast (BP: NE), Midwest (BP: MW), South (BP: South) and West (BP: West)) for which there are missing observations for the beginning of the period.[7]

Let x_{it} $(i = 1, ..., N, t = 1, ..., T)$ denote the i-th observed series at time t, where N is the total number of variables and T is the number of time series observations. When the cross-sectional dimension N is large, the use of all these variables as risk factors is impractical and even infeasible if $N \geq T$. Instead, suppose that x_{it} admits an approximate factor structure of the form

$$\begin{aligned} x_{it} &= \lambda_i' f_t + e_{it} \\ A(L) f_t &= u_t, \end{aligned} \tag{C.1}$$

where f_t is a $K \times 1$ vector of latent common factors, λ_i is a $K \times 1$ vector of latent factor loadings,

[7] We would like to thank Toni Braun for suggesting this to us and Sydney Ludvigson for making the data available on her website.

e_{it} is a vector of idiosyncratic errors, $A(L)$ is a possibly infinite dimensional lag polynomial, and u_t is a vector of *iid* errors with mean zero and a constant variance-covariance matrix. The idiosyncratic shocks are assumed to be uncorrelated with the factors at all leads and lags although serial correlation, heteroskedasticity, and a limited amount of cross-correlation is permitted (Stock and Watson, 2002a). The selection of the number of factors in the approximate factor model (C.1) is considered in Bai and Ng (2002).

Let X denote the stacked $T \times N$ data matrix with its t-th row given by $x_t' = [x_{1t}, x_{2t}, ..., x_{Nt}]$ and $F = [f_1 \ f_2 \ ... \ f_K]$ be the $T \times K$ matrix of common factors. Provided that $N, T \to \infty$, the latent factors and factor loadings can be estimated by the method of principal components by minimizing the objective function $(NT)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} (x_{it} - \lambda_i' f_t)^2$ subject to the identifying restriction $F'F/T = I_K$. Concentrating out $[\lambda_1', ..., \lambda_N']'$, the estimate of the factor matrix F, $\hat{F}$, is obtained by maximizing $\text{tr}(F'(XX')F)$ and $\hat{F}$ is a matrix of $\sqrt{T}$ times the K eigenvectors corresponding to the K largest eigenvalues of the matrix XX'. The optimal number of factors can be determined by the panel information criterion proposed by Bai and Ng (2002). While Stock and Watson (2005) and Ludvigson and Ng (2009) estimate the optimal number of factors to be 7 and 8, respectively, we restrict our attention to the first three principal components given their clear economic interpretation. More specifically, looking at the marginal R^2 from a regression of each individual series on each estimated factor, Ludvigson and Ng (2009) interpret the first estimated factor as a real activity factor, the second estimated factor as a financial factor that loads most heavily on interest rate variables, and the third estimated factor as an inflation factor. Also, note that while the factors $\hat{f}_t = [\hat{f}_{1,t}, \hat{f}_{2,t}, \hat{f}_{3,t}]'$ $(t = 1, ..., T)$ contain an estimation error, no adjustments to the standard errors of the risk premium estimates are required provided that $\sqrt{T}/N \to 0$ (Bai and Ng, 2006).

We then consider the following SDF:

$$y_t^{SW3}(\gamma) = \gamma_0 + \gamma_1 \hat{f}_{1,t} + \gamma_2 \hat{f}_{2,t} + \gamma_3 \hat{f}_{3,t}.$$

The sample period and the returns on the test assets are the same as the ones used in the monthly analysis in the paper. Table C.1 presents our results.

Table C.1 about here

The rank tests for individual factors indicate that the first (real activity) and the third (inflation) factors can be reasonably considered as useless (see Panel A). This message is further reinforced by the rank restriction test for SW3 in Panel B. Given that the second factor appears to be the only statistically significant factor (even when the misspecification-robust standard error is used), the model is re-estimated after the first and the third factors are dropped from the model. Interestingly, for the model that contains only the second estimated factor (SW1), the rank test rejects the null of reduced rank and the interest rate factor is highly statistically significant. This finding is again in line with our theoretical results and the other empirical applications.

References

Bai, Jushan, and Serena Ng, 2002, Determining the number of factors in approximate factor models, *Econometrica* 70, 191–221.

Bai, Jushan, and Serena Ng, 2006, Confidence intervals for diffusion index forecasts and inference for factor augmented regressions, *Econometrica* 74, 1133–1150.

Burnside, Craig A., 2011, The cross-section of foreign currency risk premia and consumption growth risk: Comment, *American Economic Review* 101, 3456–3476.

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

Epstein, Larry G., and Stanley E. Zin, 1989, Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework, *Econometrica* 57, 937—969.

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

Kan, Raymond, and Cesare Robotti, 2008, Specification tests of asset pricing models using excess returns, *Journal of Empirical Finance* 15, 816–838.

Liu, Laura X., and Lu Zhang, 2008, Momentum profits, factor pricing, and macroeconomic risk, *Review of Financial Studies* 21, 2417–2448.

Ludvigson, Sydney C., and Serena Ng, 2007, The empirical risk-return relation: A factor analysis approach, *Journal of Financial Economics* 83, 171–222.

Ludvigson, Sydney C., and Serena Ng, 2009, Macro factors in bond risk premia, *Review of Financial Studies* 22, 5027–5067.

Lustig, Hanno, and Adrien Verdelhan, 2007, The cross-section of foreign currency risk premia and US consumption growth risk, *American Economic Review* 97, 89–117.

Stock, James H., and Mark W. Watson, 2002a, Forecasting using principal components from a large number of predictors, *Journal of the American Statistical Association* 97, 1167–1179.

Stock, James H., and Mark W. Watson, 2002b, Macroeconomic forecasting using diffusion indexes, *Journal of Business and Economic Statistics* 20, 147–162.

Stock, James H., and Mark W. Watson, 2005, Implications of dynamic factor models for VAR analysis, NBER Working paper 11467.

Yogo, Motohiro, 2006, A consumption-based explanation of expected stock returns, *Journal of Finance* 61, 539–580.

Table A.1
Estimation and Testing Results for the Linear Models Considered by Lustig and Verdelhan (2007)

The table presents the estimation and testing results for the four linear asset pricing models considered by Lustig and Verdelhan (2007). The models considered are the consumption CAPM (CCAPM), a CCAPM specification (DCAPM) with consumption of durables and nondurables as risk factors, a linearized version of the model of Epstein and Zin (EZ-CCAPM, 1989) with consumption of nondurables and the market return as risk factors, and the durable consumption CAPM (D-CCAPM) of Yogo (2006). The models are estimated using annual excess returns on eight currency portfolios. The data are from 1953 until 2002. Panel A reports a Wald test ($\mathcal{W}$) and its p-value (p-val) of the null that $E[x_t(f_{it} - E[f_{it}])] = 0_N$. This Wald test has a χ^2_N limiting distribution. In Panel B, we report the sample HJ-distance ($\hat{\delta}$), the Lagrange multiplier (LM) test, and the rank restriction test ($\mathcal{W}^*$) with the corresponding p-values (p-val) for each model. When $K = 1$, the reported $\mathcal{W}^*$ corresponds to the $\mathcal{W}$ test described in Panel A. The t-tests under correct model specification and the model misspecification-robust t-tests 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.

Panel A: Wald Tests for Individual Factors

Test	vw	c_{nd}	c_d
$\mathcal{W}$	7.2	9.2	6.5
p-val	0.518	0.324	0.587

Panel B: HJ-Distance, Lagrange Multiplier, and Rank Tests

Model	$\hat{\delta}$	p-val	LM	p-val	$\mathcal{W}^*$	p-val
CCAPM	0.754	0.179	16.358	0.022	9.2	0.324
DCAPM	0.657	0.165	6.928	0.328	1.5	0.982
EZ-CCAPM	0.751	0.092	12.792	0.046	2.8	0.899
D-CCAPM	0.649	0.122	6.486	0.262	1.5	0.962

Panel C: t-tests Using Standard Errors Under Correct Model Specification

Model	vw	c_{nd}	c_d
CCAPM		1.90	
DCAPM		0.41	1.47
EZ-CCAPM	−0.25	2.31	
D-CCAPM	0.50	0.15	1.63

Panel D: t-tests Using Model Misspecification-Robust Standard Errors

Model	vw	c_{nd}	c_d
CCAPM		1.24	
DCAPM		0.26	0.96
EZ-CCAPM	−0.20	1.46	
D-CCAPM	0.33	0.07	0.95

Table B.1
Estimation and Testing Results for the Linear Models Considered by Liu and Zhang (2008)

The table presents the estimation and testing results for the four linear asset pricing specifications considered by Liu and Zhang (2008). The models considered are the one-factor model (MP1) with the growth rate of industrial production, *mp*, as the only risk factor, the three-factor model of Fama and French (FF3, 1993), the FF3 model augmented with *mp* (MP4), and the intertemporal CAPM-type model of Chen, Roll, and Ross (CRR, 1986). The models are estimated using monthly excess returns on ten size, ten book-to-market and ten momentum portfolios. The data are from 1960:1 until 2004:12. Panel A reports a Wald test ($\mathcal{W}$) and its p-value (p-val) of the null that $E[x_t(f_{it} - E[f_{it}])] = 0_N$. This Wald test has a χ^2_N limiting distribution. In Panel B, we report the sample HJ-distance ($\hat{\delta}$), the Lagrange multiplier (LM) test, and the rank restriction test ($\mathcal{W}^*$) with the corresponding p-values (p-val) for each model. When $K = 1$, the reported $\mathcal{W}^*$ corresponds to the $\mathcal{W}$ test described in Panel A. The t-tests under correct model specification and the model misspecification-robust t-tests 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.

Panel A: Wald Tests for Individual Factors

Test	*vw*	*smb*	*hml*	*mp*	*ui*	*dei*	*uts*	*upr*
$\mathcal{W}$	246.4	222.8	247.3	48.1	35.3	43.5	48.4	44.4
p-val	0.000	0.000	0.000	0.020	0.231	0.052	0.018	0.044

Panel B: HJ-Distance, Lagrange Multiplier, and Rank Tests

Model	$\hat{\delta}$	p-val	LM	p-val	$\mathcal{W}^*$	p-val
MP1	0.449	0.000	87.596	0.000	48.1	0.020
FF3	0.405	0.000	76.551	0.000	119.0	0.000
MP4	0.399	0.000	63.070	0.000	39.1	0.063
CRR	0.411	0.003	44.983	0.008	5.0	1.000

Panel C: t-tests Using Standard Errors Under Correct Model Specification

Model	*vw*	*smb*	*hml*	*mp*	*ui*	*dei*	*uts*	*upr*
MP1				1.87				
FF3	3.37	1.55	4.19					
MP4	3.10	1.39	3.74	1.39				
CRR				1.51	−2.25	2.24	0.91	1.25

Panel D: t-tests Using Model Misspecification-Robust Standard Errors

Model	*vw*	*smb*	*hml*	*mp*	*ui*	*dei*	*uts*	*upr*
MP1				1.10				
FF3	3.39	1.56	4.18					
MP4	3.11	1.40	3.71	0.89				
CRR				1.04	−1.15	1.00	0.45	0.59

Table C.1
Estimation and Testing Results for the Models with the Factors of Stock and Watson

The table presents the estimation and testing results for the asset pricing specifications with the factors of Stock and Watson (2002a, 2002b) constructed from 127 macroeconomic time series. SW3 denotes the model that includes the three estimated factors and SW1 denotes the model that includes only the second estimated factor. 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 ($\mathcal{W}^*$) 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 ($\mathcal{W}^*$) with the corresponding p-values (p-val) for each model. The t-tests under correct model specification and the model misspecification-robust t-tests 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.

Panel A: Rank Tests for Individual Factors

Test	$\hat{f}_1$	$\hat{f}_2$	$\hat{f}_3$
$\mathcal{W}^*$	35.7	56.8	23.2
p-val	0.076	0.000	0.564

Panel B: HJ-Distance, Lagrange Multiplier, and Rank Tests

Model	$\hat{\delta}$	p-val	LM	p-val	$\mathcal{W}^*$	p-val
SW3	0.360	0.002	49.604	0.001	18.2	0.749
SW1	0.364	0.001	56.523	0.000	56.8	0.000

Panel C: t-tests Using Standard Errors Under Correct Model Specification

Model	$\hat{f}_1$	$\hat{f}_2$	$\hat{f}_3$
SW3	0.83	−3.81	−0.31
SW1		−4.40	

Panel D: t-tests Using Model Misspecification-Robust Standard Errors

Model	$\hat{f}_1$	$\hat{f}_2$	$\hat{f}_3$
SW3	0.43	−2.59	−0.16
SW1		−3.69	