Online Appendix for

“Misspecification-Robust Inference in Linear Asset Pricing Models with Irrelevant Risk Factors”

NIKOLAY GOSPODINOV, RAYMOND KAN, and CESARE ROBOTTI

September 2013

In this online appendix, we derive the limiting distributions of the parameter estimates and their corresponding t-statistics as well as the HJ-distance test for correct model specification when a useless factor is present in the model. We follow the same notation as in the paper.

Theoretical Results for Gross Returns

First, we provide theoretical results for the gross returns case. Consider a candidate SDF which is given by

$$y_t = \tilde{f}_t'\gamma_1 + g_t\gamma_2, \tag{1}$$

where $\tilde{f}_t = [1,\ f_t']'$, f_t is a $(K-1)$-vector of useful risk factors and g_t denotes a useless factor which is independent of x_t and f_t for all time periods. For ease of exposition, we assume that $E[g_t] = 0$ and $\text{Var}[g_t] = 1$.[1] Let $B = E[x_t\tilde{f}_t']$ and note that the independence between g_t and x_t implies

$$d = E[x_t g_t] = 0_N \tag{2}$$

and

$$E[x_t x_t' g_t^2] = E[E[x_t x_t'|g_t]g_t^2] = UE[g_t^2] = U. \tag{3}$$

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

$$e(\gamma) = B\gamma_1 + d\gamma_2 - q = B\gamma_1 - q \tag{4}$$

is independent of the choice of γ_2. The pseudo-true value of the SDF parameter 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 value γ_2^* is not identified, the sample estimates of the SDF parameters are always identified and they are given by

$$\hat{\gamma} = (\hat{D}'\hat{U}^{-1}\hat{D})^{-1}\hat{D}'\hat{U}^{-1}q. \tag{5}$$

[1] This assumption 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.

Note that the estimator in (5) can be obtained equivalently by 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 construct $\hat{\gamma}_2$, 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 regress the residuals from the first projection on the residuals from the second projection. It follows that

$$\hat{\gamma}_2 = \frac{\hat{d}'\hat{U}^{-\frac{1}{2}}[I_N - \hat{U}^{-\frac{1}{2}}\hat{B}(\hat{B}'\hat{U}^{-1}\hat{B})^{-1}\hat{B}'\hat{U}^{-\frac{1}{2}}]\hat{U}^{-\frac{1}{2}}q}{\hat{d}'\hat{U}^{-\frac{1}{2}}[I_N - \hat{U}^{-\frac{1}{2}}\hat{B}(\hat{B}'\hat{U}^{-1}\hat{B})^{-1}\hat{B}'\hat{U}^{-\frac{1}{2}}]\hat{U}^{-\frac{1}{2}}\hat{d}}. \tag{6}$$

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 on the residuals from the second projection, which yields

$$\begin{aligned}\hat{\gamma}_1 &= (\hat{B}'\hat{U}^{-\frac{1}{2}}[I_N - \hat{U}^{-\frac{1}{2}}\hat{d}(\hat{d}'\hat{U}^{-1}\hat{d})^{-1}\hat{d}'\hat{U}^{-\frac{1}{2}}]\hat{U}^{-\frac{1}{2}}\hat{B})^{-1} \\ &\quad \times \hat{B}'\hat{U}^{-\frac{1}{2}}[I_N - \hat{U}^{-\frac{1}{2}}\hat{d}(\hat{d}'\hat{U}^{-1}\hat{d})^{-1}\hat{d}'\hat{U}^{-\frac{1}{2}}]\hat{U}^{-\frac{1}{2}}q. \end{aligned} \tag{7}$$

We make the following assumptions.

Assumption 1. *Assume that (i) $N > K+1$; (ii) $[x_t',\ f_t',\ g_t]'$ are jointly stationary and ergodic processes with finite fourth moments; (iii) $e_t(\gamma_1^*) - e(\gamma_1^*)$ forms a martingale difference sequence; and (iv) the matrices B $(N \times K)$ and D $(N\times(K+1))$ have a column rank K.*

Assumption 2. *Let $\epsilon_t = x_t - B(E[\tilde{f}_t\tilde{f}_t'])^{-1}\tilde{f}_t$ and assume that $E[\epsilon_t\epsilon_t'|\tilde{f}_t] = \Sigma$ (conditional homoskedasticity).*

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 $\tilde{B} = U^{-\frac{1}{2}}B$, $\tilde{q} = U^{-\frac{1}{2}}q$, 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 the conditions in Assumption 1 are satisfied.*

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

$$\sqrt{T}(\hat{\gamma}_1 - \gamma_1^*) \xrightarrow{d} (\tilde{B}'\tilde{B})^{-\frac{1}{2}}\left[r - \frac{w'u}{w'w}(\tilde{B}'\tilde{B})^{-\frac{1}{2}}\tilde{B}'z\right], \tag{8}$$

and

$$\hat{\gamma}_2 \xrightarrow{d} \frac{w'u}{w'w}. \tag{9}$$

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

$$\hat{\gamma}_1 - \gamma_1^* \xrightarrow{d} -\frac{\delta s}{w'w}(\tilde{B}'\tilde{B})^{-1}\tilde{B}'z, \tag{10}$$

and

$$\frac{1}{\sqrt{T}}\hat{\gamma}_2 \xrightarrow{d} \frac{\delta s}{w'w}. \tag{11}$$

Proof. See the Appendix.

The results in Theorem 1 subsume the results in Proposition 1 in the paper and 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 the constant zero.[2] 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 $T^{\frac{1}{2}}$.

We next derive the limiting distributions of two types of t-statistics (as defined in the paper): (i) $t_c(\hat{\gamma}_{1i})$ of $H_0 : \gamma_{1i} = \gamma_{1i}^*$ for $i = 1, \ldots, K$, and $t_c(\hat{\gamma}_2)$ of $H_0 : \gamma_2 = 0$ that use standard errors obtained under the assumption that the model is correctly specified, and (ii) $t_m(\hat{\gamma}_{1i})$ of $H_0 : \gamma_{1i} = \gamma_{1i}^*$ for $i = 1, \ldots, K$, and $t_m(\hat{\gamma}_2)$ of $H_0 : \gamma_2 = 0$ that use standard errors under potentially misspecified models. The two types of t-statistics are based on the estimated covariance matrices $\hat{\Sigma}_{\hat{\gamma}}^0 = \frac{1}{T}\sum_{t=1}^T \hat{h}_t^0 \hat{h}_t^{0\prime}$ and $\hat{\Sigma}_{\hat{\gamma}} = \frac{1}{T}\sum_{t=1}^T \hat{h}_t \hat{h}_t'$, where

$$\hat{h}_t^0 = (\hat{D}'\hat{U}^{-1}\hat{D})^{-1}\hat{D}'\hat{U}^{-1}\hat{e}_t, \tag{12}$$

$$\hat{h}_t = \hat{h}_t^0 + (\hat{D}'\hat{U}^{-1}\hat{D})^{-1}([\tilde{f}_t',\ g_t]' - \hat{D}'\hat{U}^{-1}x_t)\hat{e}'\hat{U}^{-1}x_t, \tag{13}$$

$\hat{e}_t = x_t(\tilde{f}_t'\hat{\gamma}_1 + g_t\hat{\gamma}_2) - q$ and $\hat{e} = \frac{1}{T}\sum_{t=1}^T \hat{e}_t$.

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

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

misspecified models

$$\hat{e}_t = (T^{-\frac{1}{2}}\hat{\gamma}_2)(T^{\frac{1}{2}}x_t g_t) + O_p(1) \tag{14}$$

and

$$\frac{\hat{S}}{T} = (T^{-\frac{1}{2}}\hat{\gamma}_2)^2 U + o_p(1), \tag{15}$$

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

$$\hat{S} = S + \hat{\gamma}_2^2 U + o_p(1), \tag{16}$$

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 (Proposition 2 in the paper) below provide the limiting distributions of the t-statistics under correctly specified and misspecified models.

Theorem 2.

(a) *Suppose that the conditions in Assumptions 1 and 2 hold.*[3] *If $\delta = 0$, i.e., the model is correctly specified, we have*

$$t_c(\hat{\gamma}_{1i}) \stackrel{d}{\to} \frac{\tilde{u}\tilde{z}_i + \sqrt{\lambda_i}\sqrt{w'w}\tilde{r}_i}{\left[\lambda_i w'w + \tilde{z}_i^2 + \tilde{u}^2\left(1 + \frac{\tilde{z}_i^2}{w'w}\right)\right]^{\frac{1}{2}}}, \tag{17}$$

$$t_m(\hat{\gamma}_{1i}) \stackrel{d}{\to} \frac{\tilde{u}\tilde{z}_i + \sqrt{\lambda_i}\sqrt{w'w}\tilde{r}_i}{\left[\lambda_i w'w + \tilde{z}_i^2 + \tilde{u}^2\left(1 + \frac{\tilde{z}_i^2}{w'w}\right) + \frac{\tilde{z}_i^2 v}{w'w}\right]^{\frac{1}{2}}}, \tag{18}$$

$$t_c(\hat{\gamma}_2) \stackrel{d}{\to} \frac{\tilde{u}}{\left(1 + \frac{\tilde{u}^2}{w'w}\right)^{\frac{1}{2}}}, \tag{19}$$

$$t_m(\hat{\gamma}_2) \stackrel{d}{\to} \frac{\tilde{u}}{\left(1 + \frac{\tilde{u}^2 + v}{w'w}\right)^{\frac{1}{2}}}, \tag{20}$$

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

[3]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 conditional heteroskedasticity are more involved, and the results are available upon request.

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

$$t_c(\hat{\gamma}_{1i}) \xrightarrow{d} \frac{\tilde{z}_i}{\left(1 + \frac{\tilde{z}_i^2}{w'w}\right)^{\frac{1}{2}}}, \tag{21}$$

$$t_m(\hat{\gamma}_{1i}) \xrightarrow{d} N\left(0, \frac{1}{4}\right), \tag{22}$$

$$t_c(\hat{\gamma}_2) \xrightarrow{d} \operatorname{sgn}(s)\sqrt{w'w}, \tag{23}$$

$$t_m(\hat{\gamma}_2) \xrightarrow{d} N(0,1). \tag{24}$$

Proof. See the Appendix.

Corollary 1.

(a) *Suppose that the conditions in Assumptions 1 and 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 conditions in Assumption 1 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.

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

$$\hat{\delta}^2 = \hat{e}'\hat{U}^{-1}\hat{e}. \tag{25}$$

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

$$T\hat{\delta}^2 \xrightarrow{d} \sum_{i=1}^{N-K} \xi_i X_i, \tag{26}$$

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

$$S^{\frac{1}{2}}U^{-1}S^{\frac{1}{2}} - S^{\frac{1}{2}}U^{-1}B(B'U^{-1}B)^{-1}B'U^{-1}S^{\frac{1}{2}}. \tag{27}$$

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

$$\hat{S}^{\frac{1}{2}}\hat{U}^{-1}\hat{S}^{\frac{1}{2}} - \hat{S}^{\frac{1}{2}}\hat{U}^{-1}\hat{B}(\hat{B}'\hat{U}^{-1}\hat{B})^{-1}\hat{B}'\hat{U}^{-1}\hat{S}^{\frac{1}{2}}. \tag{28}$$

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 do not hold. In the next theorem, we add to the existing literature (Kan and Zhang, 1999) 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* $\delta = 0$, *we have*

$$T\hat{\delta}^2 \xrightarrow{d} E[(\tilde{f}_t'\gamma_1^*)^2]\chi^2_{N-K-1} \tag{29}$$

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

$$\int_0^1 P\left[\chi^2_{N-K-1} > \frac{c_\alpha}{q}\right] f_{Q_1}(q)\mathrm{d}q < \alpha. \tag{30}$$

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

$$\hat{\delta}^2 \xrightarrow{d} \delta^2 Q_2 \tag{31}$$

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

$$\int_0^1 P\left[\chi^2_{N-K} > \frac{c_\alpha q}{1-q}\right] f_{Q_2}(q)\mathrm{d}q < 1. \tag{32}$$

Proof. See the Appendix.

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

Note that the limiting probabilities of rejection in (30) and (32) are only functions of the significance level α and the degree of over-identification $N-K$. Figure 1 plots these probabilities for different significance levels ($\alpha = 0.01,\ 0.05$, and 0.1) and $N-K$ ranging from 2 to 20.

Figure 1 about here

The top panel of Figure 1 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 1 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.

Theoretical Results for Excess Returns

In the following analysis, we provide theoretical results for the excess returns case. The proofs are similar to the gross returns case and are omitted but are available from the authors upon request.

Let x_t be the excess returns on N test assets at time t with mean μ and covariance matrix V. It is well known that when only excess returns are used as test assets, it is not possible to identify the mean of the candidate SDF and some normalization of the SDF becomes necessary. As a result, we follow Kan and Robotti (2008) and define the candidate SDF as

$$y_t = 1 - (f_t - \mu_f)'\gamma_1 - (g_t - \mu_g)\gamma_2, \tag{33}$$

where f_t is a vector of K systematic factors with mean μ_f and covariance matrix S_f, and g_t is a useless factor with mean μ_g and variance σ_g^2, such that it is independent of f_t and x_t for all time periods.[4]

The pseudo-true value of γ_1 under the modified HJ-distance measure is given by

$$\gamma_1^* = (B'V^{-1}B)^{-1}B'V^{-1}\mu, \tag{34}$$

[4]Note that here the number of useful factors is set equal to K. This differs from the analysis in the previous section where the number of useful factors is set equal to $K-1$.

where $B = \text{Cov}[x_t, f_t']$. We set the pseudo-true value of γ_2, γ_2^*, equal to 0 even though it is not identified (see Section 2 of the paper for a discussion of this issue). Let $d = \text{Cov}[x_t, g_t] = 0_N$, $\hat{\mu} = \frac{1}{T}\sum_{t=1}^{T} x_t$, $\hat{V} = \frac{1}{T}\sum_{t=1}^{T}(x_t - \hat{\mu})(x_t - \hat{\mu})'$, and

$$\hat{D} = \left[\frac{1}{T}\sum_{t=1}^{T} x_t(f_t - \hat{\mu}_f)',\ \frac{1}{T}\sum_{t=1}^{T} x_t(g_t - \hat{\mu}_g)\right] \equiv [\hat{B},\ \hat{d}]. \tag{35}$$

The sample estimator of $\gamma = [\gamma_1',\ \gamma_2]'$ is given by

$$\hat{\gamma} = \begin{bmatrix} \hat{\gamma}_1 \\ \hat{\gamma}_2 \end{bmatrix} = (\hat{D}'\hat{V}^{-1}\hat{D})^{-1}\hat{D}'\hat{V}^{-1}\hat{\mu}. \tag{36}$$

It is straightforward to show that

$$\begin{aligned} \hat{\gamma}_1 \quad = \quad & (\hat{B}'\hat{V}^{-\frac{1}{2}}[I_N - \hat{V}^{-\frac{1}{2}}\hat{d}(\hat{d}'\hat{V}^{-1}\hat{d})^{-1}\hat{d}'\hat{V}^{-\frac{1}{2}}]\hat{V}^{-\frac{1}{2}}\hat{B})^{-1} \\ & \times \hat{B}'\hat{V}^{-\frac{1}{2}}[I_N - \hat{V}^{-\frac{1}{2}}\hat{d}(\hat{d}'\hat{V}^{-1}\hat{d})^{-1}\hat{d}'\hat{V}^{-\frac{1}{2}}]\hat{V}^{-\frac{1}{2}}\hat{\mu} \end{aligned} \tag{37}$$

and

$$\hat{\gamma}_2 = \frac{\hat{d}'\hat{V}^{-\frac{1}{2}}[I_N - \hat{V}^{-\frac{1}{2}}\hat{B}(\hat{B}'\hat{V}^{-1}\hat{B})^{-1}\hat{B}'\hat{V}^{-\frac{1}{2}}]\hat{V}^{-\frac{1}{2}}\hat{\mu}}{\hat{d}'\hat{V}^{-\frac{1}{2}}[I_N - \hat{V}^{-\frac{1}{2}}\hat{B}(\hat{B}'\hat{V}^{-1}\hat{B})^{-1}\hat{B}'\hat{V}^{-\frac{1}{2}}]\hat{V}^{-\frac{1}{2}}\hat{d}}. \tag{38}$$

Finally, Kan and Robotti (2008) suggest that a modification of the traditional HJ-distance is needed when using the de-meaned factors. Their proposed measure, the modified HJ-distance, employs the inverse of the covariance matrix (instead of the second moment matrix) of the excess returns as the weighting matrix and is given by

$$\delta_m = \sqrt{e(\gamma_1^*)'V^{-1}e(\gamma_1^*)}, \tag{39}$$

where $e(\gamma_1^*) = \mu - B\gamma_1^*$. The sample version of the model misspecification measure in (39) is given by

$$\hat{\delta}_m = \sqrt{\hat{e}'\hat{V}^{-1}\hat{e}}, \tag{40}$$

where $\hat{e} = \hat{\mu} - \hat{D}\hat{\gamma}$.

In deriving the limiting behavior of $\hat{\gamma}_1$ and $\hat{\gamma}_2$ under correctly specified and misspecified models, we adopt the following notation. Let $\tilde{B} = V^{-\frac{1}{2}}B$, $\tilde{\mu} = V^{-\frac{1}{2}}\mu$, $e_t(\gamma_1^*) = x_t y_t^*$, $y_t^* = 1 - (f_t - \mu_f)'\gamma_1^*$, $S = E[e_t(\gamma_1^*)e_t(\gamma_1^*)']$, 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 $y \sim N(0_N, V^{-\frac{1}{2}}SV^{-\frac{1}{2}})$, and they are independent of each other. Finally, we define $w = P'z \sim N(0_{N-K}, I_{N-K})$, $s =$

$(\tilde{\mu}' P w)/(\tilde{\mu}' P P' \tilde{\mu})^{\frac{1}{2}} \sim N(0,1)$, $u = P'y \sim N(0_{N-K}, V_u)$ with $V_u = P'V^{-\frac{1}{2}} S V^{-\frac{1}{2}} P$, and $r = (\tilde{B}'\tilde{B})^{-\frac{1}{2}} \tilde{B}' y \sim N(0_K, V_r)$ with $V_r = (\tilde{B}'\tilde{B})^{-\frac{1}{2}} \tilde{B}' V^{-\frac{1}{2}} S V^{-\frac{1}{2}} \tilde{B} (\tilde{B}'\tilde{B})^{-\frac{1}{2}}$.

Theorem 4. *Assume that the conditions in Assumption 1 are satisfied.*

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

$$\sqrt{T}(\hat{\gamma}_1 - \gamma_1^*) \stackrel{d}{\to} (\tilde{B}'\tilde{B})^{-\frac{1}{2}} \left[r - \frac{w'u}{w'w} (\tilde{B}'\tilde{B})^{-\frac{1}{2}} \tilde{B}' z \right], \tag{41}$$

and

$$\hat{\gamma}_2 \stackrel{d}{\to} \frac{w'u}{\sigma_g w'w}. \tag{42}$$

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

$$\hat{\gamma}_1 - \gamma_1^* \stackrel{d}{\to} -\frac{\delta_m s}{w'w} (\tilde{B}'\tilde{B})^{-1} \tilde{B}' z, \tag{43}$$

and

$$\frac{1}{\sqrt{T}} \hat{\gamma}_2 \stackrel{d}{\to} \frac{\delta_m s}{\sigma_g w'w}. \tag{44}$$

As in the case of gross returns, 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

$$\begin{aligned} \hat{h}^0_t &= (\hat{D}'\hat{V}^{-1}\hat{D})^{-1} \hat{D}' \hat{V}^{-1} \tilde{e}_t, & (45) \\ \hat{h}_t &= \hat{h}^0_t + (\hat{D}'\hat{V}^{-1}\hat{D})^{-1} \Big([(f_t - \hat{\mu}_f)',\ g_t - \hat{\mu}_g]' - \hat{D}'\hat{V}^{-1}(x_t - \hat{\mu}) \Big) \hat{u}_t, & (46) \end{aligned}$$

$\tilde{e}_t = (x_t - \hat{\mu})\hat{y}_t + \hat{\mu}$, $\hat{y}_t = 1 - (f_t - \hat{\mu}_f)'\hat{\gamma}_1 - (g_t - \hat{\mu}_g)\hat{\gamma}_2$, and $\hat{u}_t = \hat{e}'\hat{V}^{-1}(x_t - \hat{\mu})$.

In addition to the random variables and matrices defined before Theorem 4, 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. Let c_i and $\hat{c}_i$ be the i-th diagonal elements of C and $\hat{C}$, respectively, where

$$\begin{aligned} C = \ & S_f^{-1} \mathrm{Cov}[(f_t - \mu_f)(f_t - \mu_f)', y_t^{*2}] S_f^{-1} + \gamma_1^* E[(f_t - \mu_f) y_t^{*2}]' S_f^{-1} \\ & + S_f^{-1} E[(f_t - \mu_f) y_t^{*2}] \gamma_1^{*\prime} + E[y_t^{*2}] \gamma_1^* \gamma_1^{*\prime} \end{aligned} \tag{47}$$

and

$$\hat{C} = S_f^{-1}\text{Cov}[(f_t - \mu_f)(f_t - \mu_f)', y_t^{*2}]S_f^{-1} - \gamma_1^*\gamma_1^{*\prime}. \tag{48}$$

Define

$$\lambda_i = 1 + \frac{c_i}{E[y_t^{*2}]b_i}, \tag{49}$$

$$\hat{\lambda}_i = 1 + \frac{\hat{c}_i}{E[y_t^{*2}]b_i}, \tag{50}$$

where b_i is the i-th diagonal element of $(\tilde{B}'\tilde{B})^{-1}$. Theorem 5 below provides the limiting distributions of the t-statistics under correctly specified and misspecified models. Let the following assumption replace Assumption 2.

Assumption 2'. *Let* $\epsilon_t = (x_t - \mu) - BS_f^{-1}(f_t - \mu_f)$ *and assume that* $E[\epsilon_t|f_t] = 0_N$ *and* $\text{Cov}[\epsilon_t\epsilon_t', y_t^{*2}] = 0_{N\times N}$.

Theorem 5.

(a) *Suppose that the conditions in Assumptions 1 and 2' hold. If* $\delta_m = 0$, *i.e., the model is correctly specified, we have*

$$t_c(\hat{\gamma}_{1i}) \stackrel{d}{\to} \frac{\tilde{u}\tilde{z}_i + \sqrt{\lambda_i}\sqrt{w'w}\tilde{r}_i}{\left[\hat{\lambda}_i w'w + \tilde{z}_i^2 + \tilde{u}^2\left(1 + \frac{\tilde{z}_i^2}{w'w}\right)\right]^{\frac{1}{2}}}, \tag{51}$$

$$t_m(\hat{\gamma}_{1i}) \stackrel{d}{\to} \frac{\tilde{u}\tilde{z}_i + \sqrt{\lambda_i}\sqrt{w'w}\tilde{r}_i}{\left[\hat{\lambda}_i w'w + \tilde{z}_i^2 + \tilde{u}^2\left(1 + \frac{\tilde{z}_i^2}{w'w}\right) + \frac{\tilde{z}_i^2 v}{w'w}\right]^{\frac{1}{2}}}, \tag{52}$$

$$t_c(\hat{\gamma}_2) \stackrel{d}{\to} \frac{\tilde{u}}{\left(1 + \frac{\tilde{u}^2}{w'w}\right)^{\frac{1}{2}}}, \tag{53}$$

$$t_m(\hat{\gamma}_2) \stackrel{d}{\to} \frac{\tilde{u}}{\left(1 + \frac{\tilde{u}^2 + v}{w'w}\right)^{\frac{1}{2}}}. \tag{54}$$

(b) *Suppose that the conditions in Assumption 1 hold and denote the sign operator by* $\text{sgn}(\cdot)$. *If*

$\delta_m > 0$, i.e., the model is misspecified, we have

$$t_c(\hat{\gamma}_{1i}) \xrightarrow{d} \frac{\tilde{z}_i}{\left(1 + \frac{\tilde{z}_i^2}{w'w}\right)^{\frac{1}{2}}}, \tag{55}$$

$$t_m(\hat{\gamma}_{1i}) \xrightarrow{d} N\left(0, \frac{1}{4}\right), \tag{56}$$

$$t_c(\hat{\gamma}_2) \xrightarrow{d} \mathrm{sgn}(s)\sqrt{w'w}, \tag{57}$$

$$t_m(\hat{\gamma}_2) \xrightarrow{d} N(0, 1). \tag{58}$$

In the next theorem, we characterize the limiting behavior of the sample squared modified HJ-distance in the presence of a useless factor for the excess returns case.

Theorem 6. *Let $Q_1 \sim \mathrm{Beta}\left(\frac{N-K}{2}, \frac{1}{2}\right)$ with density $f_{Q_1}(\cdot)$, $Q_2 \sim \mathrm{Beta}\left(\frac{N-K-1}{2}, \frac{1}{2}\right)$ 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 5 hold. If $\delta_m = 0$, we have*

$$T\hat{\delta}_m^2 \xrightarrow{d} E[y_t^{*2}]\chi^2_{N-K-1} \tag{59}$$

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

$$\int_0^1 P\left[\chi^2_{N-K-1} > \frac{c_\alpha}{q}\right] f_{Q_1}(q)\mathrm{d}q < \alpha. \tag{60}$$

(b) *Suppose that the assumptions in Theorem 4 hold. If $\delta_m > 0$, we have*

$$\hat{\delta}_m^2 \xrightarrow{d} \delta_m^2 Q_2 \tag{61}$$

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

$$\int_0^1 P\left[\chi^2_{N-K} > \frac{c_\alpha q}{1-q}\right] f_{Q_2}(q)\mathrm{d}q < 1. \tag{62}$$

Overall, the results for excess returns are very similar to the results for gross returns in the paper. The only noticeable differences are for the t-tests on $\hat{\gamma}_{1i}$ in part (a) of Theorem 5. This implies that the nature of the problem (and the solution) is essentially the same regardless of whether one uses gross returns or excess returns in the analysis.

Simulation Results

In this section, we undertake Monte Carlo experiments to assess the small-sample properties of the test statistics based on the modified HJ-distance in models with useful and useless factors. In addition, we analyze the finite-sample properties of some optimal GMM estimators. The simulation designs, data, and models are the same as the ones considered in Tables 1–6 of the paper.

Modified HJ-distance with excess returns

The results in Panel A of Table 1 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.

Table 1 about here

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.

Panel B of Table 1 and Table 2 present the empirical size of the t-tests in the presence of a useless factor.

Table 2 about here

The simulation results for the t-tests on the parameters of the useful factor 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. 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 5.

Tables 3 and 4 report the survival rates of different factors when using the sequential procedure described in Section 3 of the paper.

Table 3 about here

Panel A of Table 3 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 factors with nonzero SDF parameters in the model and eliminating the useless factor and the factor that does not reduce the HJ-distance. 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 factors with similarly high probability, but they produce very different results when it comes to the useless factor. For example, despite its conservative nature (due to the Bonferroni adjustment), the procedure based on t_c will retain the useless factor 30% of the time for $T = 1000$. In contrast, the procedure based on t_m will retain the useless factor only about 0.8% of the time for $T = 1000$. Similarly, the probability of at least one irrelevant factor being selected in the final specification of the model is 30% (1.5%) for $T = 1000$ when the t_c (t_m) test is used and the model is misspecified.

Table 4 about here

Table 4 reports the results from a similar exercise but this time the linear asset pricing model consists of a constant term, two useful factors with $\gamma_i^* \neq 0$ and two useless factors. This setup serves to illustrate the usefulness of combining the misspecification-robust t-tests and the Bonferroni method in controlling the false discovery rate which is about 48% (the probability that at least one useless factor is deemed priced) for the t-tests constructed under correct model specification when the true model is misspecified. In contrast, the misspecification-robust model selection procedure with the Bonferroni adjustment retains one or both useless factors only 1% of the time.

Finally, we consider a scenario in which a linear combination of two useful factors is useless.

Table 5 about here

Panel A of Table 5 shows that when the model is correctly specified, the procedures based on t_c and t_m are both effective in retaining only one factor in the model. However, when the model is misspecified (see Panel B), the procedures based on t_c and t_m deliver very different results. For $T = 1000$, the probability that both factors survive the model selection procedure based on t_c is about 38% while the probability that both factors survive the model selection procedure based on t_m is about 2%. Importantly, the probabilities that only one factor survives are very different across procedures. For example, when $T = 1000$, the probability than only one factor survives is about 89% when using t-tests under misspecified models while it is only about 56% when using t-tests under correctly specified models.

Optimal GMM with gross returns

In this subsection, we use the same notation as in the paper and set the number of useful factors equal to $K-1$. The optimal s-step ($s \geq 2$) GMM estimator of the SDF parameters is defined as

$$\hat{\gamma}^{(s)} = \left(\hat{D}'\hat{S}_{(s-1)}^{-1}\hat{D}\right)^{-1}\hat{D}'\hat{S}_{(s-1)}^{-1}q, \tag{63}$$

where

$$\hat{D} = \left[\frac{1}{T}\sum_{t=1}^{T} x_t \tilde{f}_t',\ \frac{1}{T}\sum_{t=1}^{T} x_t g_t\right] \tag{64}$$

and

$$\hat{S}_{(s-1)} = \frac{1}{T}\sum_{t=1}^{T}\left[e_t\left(\hat{\gamma}^{(s-1)}\right) - e\left(\hat{\gamma}^{(s-1)}\right)\right]\left[e_t\left(\hat{\gamma}^{(s-1)}\right) - e\left(\hat{\gamma}^{(s-1)}\right)\right]' \tag{65}$$

with $e_t\left(\hat{\gamma}^{(s-1)}\right) = x_t\left[\tilde{f}_t'\hat{\gamma}_1^{(s-1)} + g_t\hat{\gamma}_2^{(s-1)}\right] - q = x_t y_t\left(\hat{\gamma}^{(s-1)}\right) - q$, $e\left(\hat{\gamma}^{(s-1)}\right) = T^{-1}\sum_{t=1}^{T} e_t\left(\hat{\gamma}^{(s-1)}\right) = \hat{D}\hat{\gamma}^{(s-1)} - q$.

Let $\hat{u}_t = e\left(\hat{\gamma}^{(s)}\right)'\hat{S}_{(s-1)}^{-1}x_t$ and $\hat{z}_t = e\left(\hat{\gamma}^{(s)}\right)'\hat{S}_{(s-1)}^{-1}\left(e_t\left(\hat{\gamma}^{(s-1)}\right) - e\left(\hat{\gamma}^{(s-1)}\right)\right)$. A consistent estimator of the asymptotic variance of the SDF parameters under misspecified models is given by (a proof of this result is available upon request) $\hat{\Sigma}_{\hat{\gamma}^{(s)}} = \frac{1}{T}\sum\limits_{t=1}^{T}\hat{h}_t\hat{h}_t'$, where

$$\hat{h}_t = (\hat{D}'\hat{S}_{(s-1)}^{-1}\hat{D})^{-1}\left[\hat{D}'\hat{S}_{(s-1)}^{-1}\left(x_t y_t\left(\hat{\gamma}^{(s)}\right) - e_t\left(\hat{\gamma}^{(s-1)}\right)\hat{z}_t\right) + [\tilde{f}_t',\ g_t]'\hat{u}_t\right] - \hat{\gamma}^{(s)}. \tag{66}$$

When the model is correctly specified, the $\hat{h}_t$ expression simplifies to

$$\hat{h}_t^0 = \left(\hat{D}'\hat{S}_{(s-1)}^{-1}\hat{D}\right)^{-1}\hat{D}'\hat{S}_{(s-1)}^{-1}e_t\left(\hat{\gamma}^{(s)}\right). \tag{67}$$

In addition, the GMM test of correct model specification is given by

$$Te\left(\hat{\gamma}^{(s)}\right)'\hat{S}_{(s-1)}^{-1}e\left(\hat{\gamma}^{(s)}\right). \tag{68}$$

In the absence of a useless factor, it is well known that under a correctly specified model this test is asymptotically chi-squared distributed with $N-K$ degrees of freedom.

Tables 6 to 11 about here

In our simulations, we use the identity matrix to compute the first-step GMM estimator and analyze the finite-sample properties of the optimal 3-step GMM estimator and specification test in models with useful and useless factors. Our Monte Carlo simulations (see Tables 6–11) show that the results for optimal GMM are broadly consistent with the ones for the estimators and test statistics based on the HJ-distance. In addition, the rejection rates for the limiting case ($T=\infty$) are equivalent to those based on the asymptotic distributions given in Theorem 2 in the first section of this online appendix. This implies that our robust model selection procedure is also applicable to the class of optimal GMM estimators.

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*

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

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

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

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

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

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

and

$$-\sqrt{T}\hat{U}^{-\frac{1}{2}}(\hat{B}\gamma_1^* - q) \xrightarrow{d} y \sim N(0_N, V_y), \quad \text{(A.8)}$$

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

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

Therefore, we have $V_y = U^{-\frac{1}{2}} S U^{-\frac{1}{2}}$ for correctly specified models. In addition, y and z are independent of each other. Using y and z, we can write (7) as

$$\begin{aligned}
\sqrt{T}(\hat{\gamma}_1 - \gamma_1^*) &= (\hat{B}'\hat{U}^{-\frac{1}{2}}[I_N - \hat{U}^{-\frac{1}{2}}\hat{d}(\hat{d}'\hat{U}^{-1}\hat{d})^{-1}\hat{d}'\hat{U}^{-\frac{1}{2}}]\hat{U}^{-\frac{1}{2}}\hat{B})^{-1} \\
&\quad \times \hat{B}'\hat{U}^{-\frac{1}{2}}[I_N - \hat{U}^{-\frac{1}{2}}\hat{d}(\hat{d}'\hat{U}^{-1}\hat{d})^{-1}\hat{d}'\hat{U}^{-\frac{1}{2}}]\sqrt{T}\hat{U}^{-\frac{1}{2}}(q - \hat{B}\gamma_1^*) \\
&\xrightarrow{d} (\tilde{B}'[I_N - z(z'z)^{-1}z']\tilde{B})^{-1}\tilde{B}'[I_N - z(z'z)^{-1}z']y \\
&= (\tilde{B}'[I_N - z(z'z)^{-1}z']\tilde{B})^{-1}\tilde{B}'[I_N - z(z'z)^{-1}z'][PP' + \tilde{B}(\tilde{B}'\tilde{B})^{-1}\tilde{B}']y \\
&= -(\tilde{B}'[I_N - z(z'z)^{-1}z']\tilde{B})^{-1}\frac{\tilde{B}'zz'PP'y}{z'z} + (\tilde{B}'\tilde{B})^{-1}\tilde{B}'y. \quad \text{(A.10)}
\end{aligned}$$

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 identity

$$(\tilde{B}'[I_N - z(z'z)^{-1}z']\tilde{B})^{-1} = (\tilde{B}'\tilde{B})^{-1} + \frac{(\tilde{B}'\tilde{B})^{-1}\tilde{B}'zz'\tilde{B}(\tilde{B}'\tilde{B})^{-1}}{w'w} \quad \text{(A.11)}$$

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

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

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

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

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

$$\hat{\gamma}_2 = \frac{(\sqrt{T}\hat{U}^{-\frac{1}{2}}\hat{d})'(\sqrt{T}\hat{M}\hat{U}^{-\frac{1}{2}}(B - \hat{B})\gamma_1^*)}{(\sqrt{T}\hat{U}^{-\frac{1}{2}}\hat{d})'\hat{M}(\sqrt{T}\hat{U}^{-\frac{1}{2}}\hat{d})}. \quad \text{(A.14)}$$

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

$$\hat{\gamma}_2 \xrightarrow{d} \frac{z'My}{z'Mz} = \frac{(P'z)'(P'y)}{(P'z)'(P'z)} = \frac{w'u}{w'w}. \tag{A.15}$$

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 (7) as

$$\hat{\gamma}_1 \xrightarrow{d} (\tilde{B}'[I_N - z(z'^{-1}z']\tilde{B})^{-1}\tilde{B}'[I_N - z(z'^{-1}z']\tilde{q}. \tag{A.16}$$

Using (A.11) and the fact that $\gamma_1^* = (\tilde{B}'\tilde{B})^{-1}\tilde{B}'\tilde{q}$, we obtain

$$\begin{aligned}
\hat{\gamma}_1 - \gamma_1^* &\xrightarrow{d} \left[(\tilde{B}'\tilde{B})^{-1} + \frac{(\tilde{B}'\tilde{B})^{-1}\tilde{B}'zz'\tilde{B}(\tilde{B}'\tilde{B})^{-1}}{w'w}\right]\left(\tilde{B}'\tilde{q} - \frac{\tilde{B}'zz'\tilde{q}}{z'z}\right) - (\tilde{B}'\tilde{B})^{-1}\tilde{B}'\tilde{q} \\
&= -(\tilde{B}'\tilde{B})^{-1}\tilde{B}'z\frac{z'\tilde{q}}{z'z} + (\tilde{B}'\tilde{B})^{-1}\tilde{B}'z\frac{z'\tilde{B}(\tilde{B}'\tilde{B})^{-1}\tilde{B}'\tilde{q}}{w'w} - (\tilde{B}'\tilde{B})^{-1}\tilde{B}'z\frac{z'\tilde{q}}{z'z}\frac{z'\tilde{B}(\tilde{B}'\tilde{B})^{-1}\tilde{B}'z}{w'w} \\
&= -(\tilde{B}'\tilde{B})^{-1}\tilde{B}'z\frac{z'\tilde{q}}{w'w} + (\tilde{B}'\tilde{B})^{-1}\tilde{B}'z\frac{z'\tilde{B}(\tilde{B}'\tilde{B})^{-1}\tilde{B}'\tilde{q}}{w'w} \\
&= -\frac{z'M\tilde{q}}{w'w}(\tilde{B}'\tilde{B})^{-1}\tilde{B}'z \\
&= -\frac{\delta s}{w'w}(\tilde{B}'\tilde{B})^{-1}\tilde{B}'z,
\end{aligned} \tag{A.17}$$

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

$$T^{-\frac{1}{2}}\hat{\gamma}_2 = \frac{(\sqrt{T}\hat{d}'\hat{U}^{-\frac{1}{2}})\hat{M}\hat{U}^{-\frac{1}{2}}q}{(\sqrt{T}\hat{d}'\hat{U}^{-\frac{1}{2}})\hat{M}(\sqrt{T}\hat{U}^{-\frac{1}{2}}\hat{d})} \xrightarrow{d} \frac{z'M\tilde{q}}{z'Mz} = \frac{\delta s}{w'w}. \tag{A.18}$$

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

Proof of Theorem 2.

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

$$S = E[(\tilde{f}_t'\gamma_1^*)^2]U + BCB' \tag{A.19}$$

under the conditional homoskedasticity assumption. It follows that

$$\begin{aligned}
V_u &= P'^{-\frac{1}{2}}SU^{-\frac{1}{2}}P = E[(\tilde{f}_t'\gamma_1^*)^2]I_{N-K}, & \text{(A.20)} \\
V_r &= (\tilde{B}'\tilde{B})^{-\frac{1}{2}}\tilde{B}'^{-\frac{1}{2}}SU^{-\frac{1}{2}}\tilde{B}(\tilde{B}'\tilde{B})^{-\frac{1}{2}} = E[(\tilde{f}_t'\gamma_1^*)^2]I_K + (\tilde{B}'\tilde{B})^{\frac{1}{2}}C(\tilde{B}'\tilde{B})^{\frac{1}{2}}, & \text{(A.21)} \\
\mathrm{Cov}[u, r'] &= P'^{-\frac{1}{2}}SU^{-\frac{1}{2}}\tilde{B}(\tilde{B}'\tilde{B})^{-\frac{1}{2}} = 0_{(N-K)\times K}. & \text{(A.22)}
\end{aligned}$$

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.12) to

$$\sqrt{T}(\hat{\gamma}_1 - \gamma_1^*) \xrightarrow{d} -E[(\tilde{f}_t'\gamma_1^*)^2]^{\frac{1}{2}} \frac{\tilde{u}}{(w'w)^{\frac{1}{2}}} (\tilde{B}'\tilde{B})^{-1}\tilde{B}'z + (\tilde{B}'\tilde{B})^{-\frac{1}{2}} r. \tag{A.23}$$

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

$$\hat{V}_m(\hat{\gamma}) = \frac{1}{T^2} \sum_{t=1}^{T} \hat{h}_t \hat{h}_t', \tag{A.24}$$

where

$$\hat{h}_t = (\hat{D}'\hat{U}^{-1}\hat{D})^{-1}\hat{D}'\hat{U}^{-1}\hat{e}_t + (\hat{D}'\hat{U}^{-1}\hat{D})^{-1}([\tilde{f}_t',\ g_t]' - \hat{D}'\hat{U}^{-1}x_t)\hat{u}_t, \tag{A.25}$$

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

It is straightforward to show that

$$\hat{D}'\hat{U}^{-1} = \begin{bmatrix} \tilde{B}'U^{-\frac{1}{2}} + O_p(T^{-\frac{1}{2}}) \\ \frac{1}{\sqrt{T}} z'U^{-\frac{1}{2}} + O_p(T^{-1}) \end{bmatrix}, \tag{A.26}$$

$$\hat{D}'\hat{U}^{-1}\hat{D} = \begin{bmatrix} \tilde{B}'\tilde{B} + O_p(T^{-\frac{1}{2}}) & \frac{1}{\sqrt{T}}\tilde{B}'z + O_p(T^{-1}) \\ \frac{1}{\sqrt{T}} z'\tilde{B} + O_p(T^{-1}) & \frac{z'z}{T} + O_p(T^{-\frac{3}{2}}) \end{bmatrix}. \tag{A.27}$$

Then, using the partitioned matrix inverse formula, we have

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

where

$$H = (\tilde{B}'[I_N - z(z'z)^{-1}z']\tilde{B})^{-1} = (\tilde{B}'\tilde{B})^{-1} + \frac{(\tilde{B}'\tilde{B})^{-1}\tilde{B}'zz'\tilde{B}(\tilde{B}'\tilde{B})^{-1}}{w'w}. \tag{A.29}$$

After simplification, we obtain

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

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

$$\hat{e} = \hat{D}\hat{\gamma} - q = \hat{D}(\hat{D}'\hat{U}^{-1}\hat{D})^{-1}\hat{D}'\hat{U}^{-1}q - q. \tag{A.31}$$

Using (A.13), (A.15), and the identity

$$I_N - \hat{U}^{-\frac{1}{2}}\hat{D}(\hat{D}'\hat{U}^{-1}\hat{D})^{-1}\hat{D}'\hat{U}^{-\frac{1}{2}} = \hat{M} - \hat{M}\hat{U}^{-\frac{1}{2}}\hat{d}(\hat{d}'\hat{U}^{-\frac{1}{2}}\hat{M}\hat{U}^{-\frac{1}{2}}\hat{d})^{-1}\hat{d}'\hat{U}^{-\frac{1}{2}}\hat{M}, \tag{A.32}$$

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

$$-\sqrt{T}\hat{U}^{-\frac{1}{2}}\hat{e} = \sqrt{T}\hat{M}\hat{U}^{-\frac{1}{2}}q - \sqrt{T}\hat{M}\hat{U}^{-\frac{1}{2}}\hat{d}\hat{\gamma}_2 \stackrel{d}{\to} My - Mz\frac{w'u}{w'w} = P\left(I_{N-K} - \frac{ww'}{w'w}\right)u, \tag{A.33}$$

and we have

$$\sqrt{T}\hat{u}_t \stackrel{d}{\to} -u'\left(I_{N-K} - \frac{ww'}{w'w}\right)P'^{-\frac{1}{2}}x_t. \tag{A.34}$$

Using (A.28), (A.30), (A.34), and the fact that

$$\hat{e}_t = x_t(\tilde{f}_t'\hat{\gamma}_1 + \hat{\gamma}_2 g_t) - q = x_t\tilde{f}_t'\gamma_1^* - q + \frac{w'u}{w'w}x_t g_t + O_p(T^{-\frac{1}{2}}) \tag{A.35}$$

under a correctly specified model, we can write the limiting distribution of $\hat{h}_t = [\hat{h}_{1t}',\ \hat{h}_{2t}]'$, where $\hat{h}_{1t}$ denotes the first K elements of $\hat{h}_t$, as

$$\begin{aligned}\hat{h}_{1t} \quad \stackrel{d}{\to} \quad & \left[(\tilde{B}'\tilde{B})^{-1}\tilde{B}'U^{-\frac{1}{2}} - \frac{(\tilde{B}'\tilde{B})^{-1}\tilde{B}'zw'P'U^{-\frac{1}{2}}}{w'w}\right]\left(x_t\tilde{f}_t'\gamma_1^* - q + x_t g_t\frac{w'u}{w'w}\right) \\ & + \frac{(\tilde{B}'\tilde{B})^{-1}\tilde{B}'z}{w'w}u'\left(I_{N-K} - \frac{ww'}{w'w}\right)P'^{-\frac{1}{2}}x_t g_t,\end{aligned} \tag{A.36}$$

$$\frac{\hat{h}_{2t}}{\sqrt{T}} \stackrel{d}{\to} \frac{1}{w'w}w'P'U^{-\frac{1}{2}}\left(x_t\tilde{f}_t'\gamma_1^* - q + x_t g_t\frac{w'u}{w'w}\right) - \frac{1}{w'w}u'\left(I_{N-K} - \frac{ww'}{w'w}\right)P'U^{-\frac{1}{2}}x_t g_t. \tag{A.37}$$

Under the conditional homoskedasticity assumption, we have

$$\frac{1}{T}\sum_{t=1}^{T}(x_t\tilde{f}_t'\gamma_1^* - q)(x_t\tilde{f}_t'\gamma_1^* - q)' \stackrel{\text{a.s.}}{\longrightarrow} S = E[(\tilde{f}_t'\gamma_1^*)^2]U + BCB'. \tag{A.38}$$

Together with the fact that

$$\frac{1}{T}\sum_{t=1}^{T}x_t x_t' g_t^2 \stackrel{\text{a.s.}}{\longrightarrow} E[x_t x_t' g_t^2] = E[x_t x_t']E[g_t^2] = U, \tag{A.39}$$

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

$$\begin{aligned}T\hat{V}_m(\hat{\gamma}_1) \quad &= \quad \frac{1}{T}\sum_{t=1}^{T}\hat{h}_{1t}\hat{h}_{1t}' \\ &\stackrel{d}{\to} \quad E[(\tilde{f}_t'\gamma_1^*)^2]\left(1 + \frac{\tilde{u}^2}{w'w}\right)\left[(\tilde{B}'\tilde{B})^{-1} + \frac{(\tilde{B}'\tilde{B})^{-1}\tilde{B}'zz'\tilde{B}(\tilde{B}'\tilde{B})^{-1}}{w'w}\right] + C \\ &\qquad + u'\left(I_{N-K} - \frac{ww'}{w'w}\right)u\frac{(\tilde{B}'\tilde{B})^{-1}\tilde{B}'zz'\tilde{B}(\tilde{B}'\tilde{B})^{-1}}{(w'w)^2}.\end{aligned} \tag{A.40}$$

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

$$\tilde{z}_i = -\frac{\boldsymbol{\iota}_i'(\tilde{B}'\tilde{B})^{-1}\tilde{B}'z}{\sqrt{b_i}} \sim N(0,1), \tag{A.41}$$

$$v = \frac{u'[I_{N-K} - w(w'w)^{-1}w']u}{E[(\tilde{f}_t'\gamma_1^*)^2]} \sim \chi^2_{N-K-1}, \tag{A.42}$$

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

$$Ts_m^2(\hat{\gamma}_{1i}) = T\boldsymbol{\iota}_i'\hat{V}_m(\hat{\gamma}_1)\boldsymbol{\iota}_i \xrightarrow{d} E[(\tilde{f}_t'\gamma_1^*)^2]b_i\left[\left(1+\frac{\tilde{u}^2}{w'w}\right)\left(1+\frac{\tilde{z}_i^2}{w'w}\right)+\frac{\tilde{z}_i^2 v}{(w'w)^2}\right]+c_i, \tag{A.43}$$

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

$$\tilde{r}_i = (E[(\tilde{f}_t'\gamma_1^*)^2]b_i + c_i)^{-\frac{1}{2}}\boldsymbol{\iota}_i'(\tilde{B}'\tilde{B})^{-\frac{1}{2}}r \sim N(0,1), \tag{A.44}$$

we can write the i-th element in (A.23) as

$$\sqrt{T}(\hat{\gamma}_{1i}-\gamma_{1i}^*) \xrightarrow{d} (E[(\tilde{f}_t'\gamma_1^*)^2]b_i)^{\frac{1}{2}}\frac{\tilde{u}\tilde{z}_i}{(w'w)^{\frac{1}{2}}} + (E[(\tilde{f}_t'\gamma_1^*)^2]b_i + c_i)^{\frac{1}{2}}\tilde{r}_i. \tag{A.45}$$

Finally, by letting[5]

$$\lambda_i = 1 + \frac{c_i}{E[(\tilde{f}_t'\gamma_1^*)^2]b_i} > 0, \tag{A.46}$$

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

$$t_m(\hat{\gamma}_{1i}) = \frac{\hat{\gamma}_{1i}-\gamma_{1i}^*}{s_m(\hat{\gamma}_{1i})} \xrightarrow{d} \frac{\tilde{u}\tilde{z}_i + \sqrt{\lambda_i}\sqrt{w'w}\tilde{r}_i}{\left[\lambda_i(w'w)+\tilde{z}_i^2+\tilde{u}^2\left(1+\frac{\tilde{z}_i^2}{w'w}\right)+\frac{\tilde{z}_i^2 v}{w'w}\right]^{\frac{1}{2}}}. \tag{A.47}$$

The estimated covariance matrix of $\hat{\gamma}_1$ that assumes a correctly specified model is obtained by dropping the second term in (A.40). Then, it can be shown that

$$Ts_c^2(\hat{\gamma}_{1i}) \xrightarrow{d} E[(\tilde{f}_t'\gamma_1^*)^2]b_i\left[\left(1+\frac{\tilde{u}^2}{w'w}\right)\left(1+\frac{\tilde{z}_i^2}{w'w}\right)\right]+c_i \tag{A.48}$$

and hence

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

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

$$\hat{\gamma}_2 \xrightarrow{d} \frac{w'u}{w'w} = \frac{(w'V_u w)^{\frac{1}{2}}}{(w'w)}\tilde{u}, \tag{A.50}$$

[5]From (A.44), 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 $\lambda_i > 0$.

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

$$\begin{aligned} s_m^2(\hat{\gamma}_2) &= \frac{1}{T^2}\sum_{t=1}^T \hat{h}_{2t}^2 \\ &\stackrel{d}{\to} \frac{1}{(w'w)^2}\left[w'V_u w + \frac{(w'u)^2}{w'w}\right] + \frac{u'[I_{N-K} - w(w'w)^{-1}w']u}{(w'w)^2} \\ &= \frac{w'V_u w + u'u}{(w'w)^2}. \end{aligned} \tag{A.51}$$

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

$$t_m(\hat{\gamma}_2) = \frac{\hat{\gamma}_2}{s_m(\hat{\gamma}_2)} \stackrel{d}{\to} \frac{\tilde{u}}{\left(1 + \frac{u'u}{w'V_u w}\right)^{\frac{1}{2}}}. \tag{A.52}$$

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

$$s_c^2(\hat{\gamma}_2) \stackrel{d}{\to} \frac{1}{(w'w)^2}\left[w'V_u w + \frac{(w'u)^2}{w'w}\right] = \frac{w'V_u w}{(w'w)^2}\left(1 + \frac{\tilde{u}^2}{w'w}\right). \tag{A.53}$$

It follows that

$$t_c(\hat{\gamma}_2) = \frac{\hat{\gamma}_2}{s_c(\hat{\gamma}_2)} \stackrel{d}{\to} \frac{\tilde{u}}{\left(1 + \frac{\tilde{u}^2}{w'w}\right)^{\frac{1}{2}}}. \tag{A.54}$$

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

$$t_m(\hat{\gamma}_2) \stackrel{d}{\to} \frac{\tilde{u}}{\left(1 + \frac{\tilde{u}^2 + v}{w'w}\right)^{\frac{1}{2}}}, \tag{A.55}$$

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

part (b): We first derive the limiting distribution of $\hat{h}_t$ in (A.25). When a model is misspecified, we can see from part (b) of Theorem 1 that $\hat{\gamma}_2 = O_p(T^{\frac{1}{2}})$ and $\hat{\gamma}_1 = O_p(1)$, so $\hat{\gamma}_2$ is the dominant term. Therefore, using (11), we have

$$\hat{e}_t = x_t(\tilde{f}_t'\hat{\gamma}_1 + g_t\hat{\gamma}_2) - q = x_t g_t \hat{\gamma}_2 + O_p(1) = \frac{\sqrt{T}\delta s}{w'w} x_t g_t + O_p(1). \tag{A.56}$$

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

$$-\hat{U}^{-\frac{1}{2}}\hat{e} = \hat{M}\hat{U}^{-\frac{1}{2}}q - \hat{M}\hat{U}^{-\frac{1}{2}}\hat{d}\hat{\gamma}_2 \stackrel{d}{\to} M\tilde{q} - \frac{Mzz'M\tilde{q}}{z'Mz} = P[I_{N-K} - w(w'w)^{-1}w']P'\tilde{q}. \tag{A.57}$$

It follows that under a misspecified model,

$$\hat{u}_t = \hat{e}'\hat{U}^{-1}x_t \stackrel{d}{\to} -\tilde{q}'P[I_{N-K} - w(w'w)^{-1}w']P'^{-\frac{1}{2}}x_t. \tag{A.58}$$

Then, using (A.28) and (A.30), we can express the limiting distribution of $\hat{h}_t = [\hat{h}_{1t}',\ \hat{h}_{2t}]'$ as

$$
\begin{aligned}
\frac{\hat{h}_{1t}}{\sqrt{T}} &\xrightarrow{d} \frac{\tilde{q}'Pw}{w'w}(\tilde{B}'\tilde{B})^{-1}\tilde{B}'\left(I_N - \frac{zw'}{w'w}P'\right)U^{-\frac{1}{2}}x_t g_t \\
&\quad + \frac{(\tilde{B}'\tilde{B})^{-1}(\tilde{B}'z)}{w'w}\tilde{q}'P[I_{N-K} - w(w'w)^{-1}w']P'^{-\frac{1}{2}}x_t g_t, \qquad \text{(A.59)} \\
\frac{\hat{h}_{2t}}{T} &\xrightarrow{d} \frac{\tilde{q}'Pw}{(w'w)^2}w'P'^{-\frac{1}{2}}x_t g_t - \frac{1}{w'w}\tilde{q}'P[I_{N-K} - w(w'w)^{-1}w']P'^{-\frac{1}{2}}x_t g_t. \qquad \text{(A.60)}
\end{aligned}
$$

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

$$
\tilde{B}'\left(I_N - \frac{zw'}{w'w}P'\right)P[I_{N-K} - w(w'w)^{-1}w']P'\tilde{q} = 0_K, \qquad \text{(A.61)}
$$

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

$$
\begin{aligned}
\hat{V}_m(\hat{\gamma}_1) &= \frac{1}{T^2}\sum_{t=1}^{T}\hat{h}_{1t}\hat{h}_{1t}' \\
&= \frac{(\tilde{q}'Pw)^2}{(w'w)^2}\left[(\tilde{B}'\tilde{B})^{-1} + \frac{(\tilde{B}'\tilde{B})^{-1}\tilde{B}'zz'\tilde{B}(\tilde{B}'\tilde{B})^{-1}}{w'w}\right] \\
&\quad + \frac{1}{(w'w)^2}\left[\tilde{q}'PP'\tilde{q} - \frac{(\tilde{q}'Pw)^2}{w'w}\right](\tilde{B}'\tilde{B})^{-1}\tilde{B}'zz'\tilde{B}(\tilde{B}'\tilde{B})^{-1} \\
&= \frac{\delta^2}{(w'w)^2}\left[s^2(\tilde{B}'\tilde{B})^{-1} + (\tilde{B}'\tilde{B})^{-1}\tilde{B}'zz'\tilde{B}(\tilde{B}'\tilde{B})^{-1}\right]. \qquad \text{(A.62)}
\end{aligned}
$$

Using $\tilde{z}_i$ as defined in (A.41), we can express the limiting distribution of $s_m^2(\hat{\gamma}_{1i})$ as

$$
s_m^2(\hat{\gamma}_{1i}) = \boldsymbol{\iota}_i'\hat{V}_m(\hat{\gamma}_1)\boldsymbol{\iota}_i \xrightarrow{d} \frac{\delta^2 b_i}{(w'w)^2}(s^2 + \tilde{z}_i^2). \qquad \text{(A.63)}
$$

In addition, we can also use $\tilde{z}_i$ to express the i-th element in (10) as

$$
\hat{\gamma}_{1i} - \gamma_{1i}^* \xrightarrow{d} \frac{\delta s\sqrt{b_i}\tilde{z}_i}{w'w}. \qquad \text{(A.64)}
$$

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

$$
t_m(\hat{\gamma}_{1i}) = \frac{\hat{\gamma}_{1i} - \gamma_{1i}^*}{s_m(\hat{\gamma}_{1i})} \xrightarrow{d} \frac{s\tilde{z}_i}{\sqrt{s^2 + \tilde{z}_i^2}}. \qquad \text{(A.65)}
$$

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

$$
f(\omega, \theta) = \frac{\omega e^{-\frac{\omega^2}{2}}}{2\pi}I_{\{\omega>0\}}I_{\{0<\theta<2\pi\}}. \qquad \text{(A.66)}
$$

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

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

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

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

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.62). We can then show that

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

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

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

Turning our attention to the limiting distributions of $t_c(\hat{\gamma}_2)$ and $t_m(\hat{\gamma}_2)$, we use (A.60) and the fact that $\delta^2 = \tilde{q}'PP'\tilde{q}$ to obtain

$$\begin{aligned} \frac{s_m^2(\hat{\gamma}_2)}{T} &= \frac{1}{T^3}\sum_{t=1}^{T}\hat{h}_{2t}^2 \\ &\stackrel{d}{\to} \frac{(\tilde{q}'^2}{(w'w)^4}w'w + \frac{1}{(w'w)^2}\tilde{q}'P\left(I_{N-K}-\frac{ww'}{w'w}\right)P'\tilde{q} \\ &= \frac{\delta^2}{(w'w)^2}. \end{aligned} \tag{A.71}$$

Therefore, using (11), the t-statistic of $\hat{\gamma}_2$ under the misspecification-robust standard error is given by

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

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

$$\frac{s_c^2(\hat{\gamma}_2)}{T} \stackrel{d}{\to} \frac{(\tilde{q}'Pw)^2}{(w'w)^3} = \frac{\delta^2 s^2}{(w'w)^3}. \tag{A.73}$$

It follows that

$$t_c(\hat{\gamma}_2) = \frac{\hat{\gamma}_2}{s_c(\hat{\gamma}_2)} \xrightarrow{d} \operatorname{sgn}(s)\sqrt{w'w}. \tag{A.74}$$

Note that since $s \sim N(0,1)$, $\operatorname{sgn}(s)$ has probabilities of $1/2$ of taking the values of -1 or 1, and it is independent of s^2. As a result, $\operatorname{sgn}(s)$ is also independent of $w'w \sim \chi^2_{N-K}$.[6] This completes the proof of part (b) of Theorem 2.

Proof of Corollary 1 (Proposition 2 in the paper).

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

$$t_c^2(\hat{\gamma}_{1i}) \xrightarrow{d} \frac{(\tilde{u}\tilde{z}_i + \sqrt{\lambda_i}\sqrt{w'w}\tilde{r}_i)^2}{\lambda_i(w'w) + \tilde{z}_i^2 + \tilde{u}^2\left(1 + \frac{\tilde{z}_i^2}{w'w}\right)}. \tag{A.76}$$

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.33) in the proof of Theorem 2, we can easily obtain

$$T\hat{\delta}^2 = T\hat{e}'\hat{U}^{-1}\hat{e} \xrightarrow{d} u'[I_{N-K} - w(w'w)^{-1}w']u = u'P_w P_w' u, \tag{A.77}$$

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

$$f(t) = \frac{|t|^{N-K-1}e^{-\frac{t^2}{2}}}{2^{\frac{N-K}{2}}\Gamma\left(\frac{N-K}{2}\right)}. \tag{A.75}$$

where P_w is an $(N-K)\times(N-K-1)$ orthonormal matrix such that $P_wP_w' = I_{N-K} - w(w'w)^{-1}w'$. Let $\tilde{v} = (P_w'V_uP_w)^{-\frac{1}{2}}P_w'u \sim N(0_{N-K-1}, I_{N-K-1})$, which is independent of w. Then, we have

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

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

$$\hat{S}^{\frac{1}{2}}\hat{U}^{-1}\hat{S}^{\frac{1}{2}} - \hat{S}^{\frac{1}{2}}\hat{U}^{-1}\hat{D}(\hat{D}'\hat{U}^{-1}\hat{D})^{-1}\hat{D}'\hat{U}^{-1}\hat{S}^{\frac{1}{2}}. \tag{A.79}$$

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

$$\begin{aligned} & \hat{S}^{\frac{1}{2}}\hat{U}^{-\frac{1}{2}}[I_N - \hat{U}^{-\frac{1}{2}}\hat{D}(\hat{D}'\hat{U}^{-1}\hat{D})^{-1}\hat{D}'\hat{U}^{-\frac{1}{2}}]\hat{U}^{-\frac{1}{2}}\hat{S}^{\frac{1}{2}} \\ = \ & \hat{S}^{\frac{1}{2}}\hat{U}^{-\frac{1}{2}}\hat{M}\hat{U}^{-\frac{1}{2}}\hat{S}^{\frac{1}{2}} - \hat{S}^{\frac{1}{2}}\hat{U}^{-\frac{1}{2}}\hat{M}\hat{U}^{-\frac{1}{2}}\hat{d}(\hat{d}'\hat{U}^{-\frac{1}{2}}\hat{M}\hat{U}^{-\frac{1}{2}}\hat{d})^{-1}\hat{d}'\hat{U}^{-\frac{1}{2}}\hat{M}\hat{U}^{-\frac{1}{2}}\hat{S}^{\frac{1}{2}}. \end{aligned} \tag{A.80}$$

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

$$\hat{S}^{\frac{1}{2}}\hat{U}^{-\frac{1}{2}}\hat{P}\hat{P}_w\hat{P}_w'\hat{P}'\hat{U}^{-\frac{1}{2}}\hat{S}^{\frac{1}{2}}, \tag{A.81}$$

or equivalently the eigenvalues of

$$\hat{P}_w'\hat{P}'\hat{U}^{-\frac{1}{2}}\hat{S}\hat{U}^{-\frac{1}{2}}\hat{P}\hat{P}_w. \tag{A.82}$$

Using (A.35), we can show that

$$\hat{P}'\hat{U}^{-\frac{1}{2}}\hat{e}_t \xrightarrow{d} P'^{-\frac{1}{2}}e_t(\gamma_1^*) + \frac{w'u}{w'w}P'^{-\frac{1}{2}}x_tg_t. \tag{A.83}$$

It follows that

$$\hat{P}'\hat{U}^{-\frac{1}{2}}\hat{S}\hat{U}^{-\frac{1}{2}}\hat{P} \xrightarrow{d} P'U^{-\frac{1}{2}}SU^{-\frac{1}{2}}P + \frac{(w'u)^2}{(w'w)^2}I_{N-K} = V_u + \frac{(w'V_uw)\tilde{u}^2}{(w'w)^2}I_{N-K}, \tag{A.84}$$

where $\tilde{u} = w'u/(w'V_uw)^{\frac{1}{2}} \sim N(0,1)$ 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

$$\begin{aligned} T\hat{\delta}^2 & \xrightarrow{d} E[(\tilde{f}_t'\gamma_1^*)^2]\tilde{v}'\tilde{v} \sim E[(\tilde{f}_t'\gamma_1^*)^2]\chi^2_{N-K-1}, && \text{(A.85)} \\ \hat{P}_w'\hat{P}'\hat{U}^{-\frac{1}{2}}\hat{S}\hat{U}^{-\frac{1}{2}}\hat{P}\hat{P}_w & \xrightarrow{d} E[(\tilde{f}_t'\gamma_1^*)^2]\left(1 + \frac{\tilde{u}^2}{w'w}\right)I_{N-K-1}. && \text{(A.86)} \end{aligned}$$

It follows that

$$\hat{\xi}_i \stackrel{d}{\to} E[(\tilde{f}_t'\gamma_1^*)^2]\left(1+\frac{\tilde{u}^2}{w'w}\right)=\frac{E[(\tilde{f}_t'\gamma_1^*)^2]}{Q_1}, \tag{A.87}$$

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

$$\int_0^1 P\left[\chi^2_{N-K-1} > \frac{c_\alpha}{q}\right] f_{Q_1}(q)\mathrm{d}q, \tag{A.88}$$

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.57), the limiting distribution of the squared sample HJ-distance $\hat{\delta}^2 = \hat{e}'\hat{U}^{-1}\hat{e}$ can be obtained as

$$\begin{aligned}\hat{\delta}^2 &\stackrel{d}{\to} \tilde{q}'P[I_{N-K} - w(w'w)^{-1}w']P'\tilde{q} \\ &= (\tilde{q}'PP'\tilde{q})\frac{w'[I_{N-K}-P'\tilde{q}(\tilde{q}'PP'\tilde{q})^{-1}\tilde{q}'P]w}{w'w} = \delta^2 Q_2,\end{aligned} \tag{A.89}$$

where

$$Q_2 = \frac{w'[I_{N-K}-P'\tilde{q}(\tilde{q}'PP'\tilde{q})^{-1}\tilde{q}'P]w}{w'w} \sim \text{Beta}\left(\frac{N-K-1}{2},\frac{1}{2}\right) \tag{A.90}$$

and it is independent of w.

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

$$\hat{P}_w'\hat{P}'\hat{U}^{-\frac{1}{2}}\hat{S}\hat{U}^{-\frac{1}{2}}\hat{P}\hat{P}_w. \tag{A.91}$$

From (15) and (11), we have

$$\frac{\hat{S}}{T} \stackrel{d}{\to} \frac{\delta^2 s^2}{(w'w)^2}U, \tag{A.92}$$

which implies

$$\frac{\hat{P}_w'\hat{P}'\hat{U}^{-\frac{1}{2}}\hat{S}\hat{U}^{-\frac{1}{2}}\hat{P}\hat{P}_w}{T} \stackrel{d}{\to} \frac{\delta^2 s^2}{(w'w)^2}I_{N-K-1} \tag{A.93}$$

and

$$\frac{\hat{\xi}_i}{T} \stackrel{d}{\to} \frac{\delta^2 s^2}{(w'w)^2} = \frac{\delta^2(1-Q_2)}{w'w}. \tag{A.94}$$

When we compare $T\hat{\delta}^2$ with the distribution of $\sum_{i=1}^{N-K-1}\hat{\xi}_i X_i$, we are effectively comparing Q_2 with $(1-Q_2)/(w'^2_{N-K-1}$, and we will reject $H_0: \delta = 0$ when

$$w'w > \frac{c_\alpha Q_2}{1-Q_2}. \tag{A.95}$$

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

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

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

References

Hansen, L. P., J. Heaton, and E. G. J. Luttmer, 1995, "Econometric Evaluation of Asset Pricing Models," *Review of Financial Studies*, 8, 237–274.

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

Kan, R., and C. Robotti, 2008, "Specification Tests of Asset Pricing Models Using Excess Returns," *Journal of Empirical Finance*, 15, 816–838.

Kan, R., and C. Zhang, 1999, "GMM Tests of Stochastic Discount Factor Models with Useless Factors," *Journal of Financial Economics*, 54, 103–127.

Table 1
Empirical size of the t-tests (modified HJ-distance case)

Panel A: Model with a useful factor

		Correctly specified model			Misspecified model		
t-test	T	10%	5%	1%	10%	5%	1%
t_c	200	0.098	0.049	0.009	0.098	0.049	0.009
	600	0.100	0.050	0.009	0.099	0.048	0.009
	1000	0.097	0.048	0.010	0.099	0.049	0.009
	∞	0.100	0.050	0.010	0.100	0.050	0.010
t_m	200	0.098	0.049	0.009	0.098	0.048	0.009
	600	0.100	0.050	0.009	0.098	0.048	0.009
	1000	0.097	0.048	0.010	0.099	0.049	0.009
	∞	0.100	0.050	0.010	0.100	0.050	0.010

Panel B: Model with a useless factor

		Correctly specified model			Misspecified model		
t-test	T	10%	5%	1%	10%	5%	1%
t_c	200	0.129	0.067	0.013	0.327	0.235	0.101
	600	0.101	0.046	0.007	0.472	0.384	0.231
	1000	0.095	0.044	0.006	0.556	0.477	0.328
	∞	0.088	0.039	0.005	1.000	1.000	1.000
t_m	200	0.037	0.012	0.001	0.080	0.036	0.005
	600	0.022	0.006	0.000	0.082	0.038	0.006
	1000	0.021	0.006	0.000	0.088	0.041	0.007
	∞	0.018	0.004	0.000	0.100	0.050	0.010

The table presents the empirical size of the t-tests of $H_0 : \gamma_1 = \gamma_1^*$ in a model with a useful factor (Panel A) and in a model with a useless factor (Panel B). Each panel considers the case in which the model is correctly specified and the case in which the model is misspecified. 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 excess returns on the 25 Fama-French size and book-to-market portfolios and the 17 Fama-French industry portfolios for the period 1959:2–2012:12. The various t-statistics are compared to the critical values from a standard normal distribution. In Panel B, the rejection rates for the limiting case ($T = \infty$) are based on the asymptotic distributions given in Theorem 5.

Table 2
Empirical size of the *t*-tests (modified HJ-distance case)

Panel A: Correctly specified model

t-test	T	$\hat{\gamma}_1$ 10%	$\hat{\gamma}_1$ 5%	$\hat{\gamma}_1$ 1%	$\hat{\gamma}_2$ 10%	$\hat{\gamma}_2$ 5%	$\hat{\gamma}_2$ 1%
t_c	200	0.094	0.045	0.008	0.130	0.066	0.012
	600	0.095	0.047	0.009	0.100	0.047	0.007
	1000	0.097	0.048	0.009	0.095	0.043	0.006
	∞	0.092	0.045	0.008	0.088	0.039	0.005
t_m	200	0.090	0.042	0.008	0.036	0.012	0.001
	600	0.091	0.044	0.008	0.023	0.006	0.000
	1000	0.093	0.046	0.008	0.020	0.005	0.000
	∞	0.088	0.042	0.008	0.018	0.004	0.000

Panel B: Misspecified model

t-test	T	$\hat{\gamma}_1$ 10%	$\hat{\gamma}_1$ 5%	$\hat{\gamma}_1$ 1%	$\hat{\gamma}_2$ 10%	$\hat{\gamma}_2$ 5%	$\hat{\gamma}_2$ 1%
t_c	200	0.094	0.046	0.008	0.321	0.230	0.098
	600	0.095	0.047	0.008	0.464	0.374	0.223
	1000	0.094	0.046	0.008	0.553	0.471	0.321
	∞	0.088	0.039	0.005	1.000	1.000	1.000
t_m	200	0.086	0.041	0.007	0.080	0.036	0.005
	600	0.079	0.036	0.006	0.081	0.038	0.006
	1000	0.072	0.032	0.005	0.088	0.041	0.007
	∞	0.001	0.000	0.000	0.100	0.050	0.010

The table presents the empirical size of the t-tests of $H_0 : \gamma_i = \gamma_i^*$ $(i = 1, 2)$ in a model with a useful and a useless factor. 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 excess returns on the 25 Fama-French size and book-to-market portfolios and the 17 Fama-French industry portfolios for the period 1959:2–2012:12. The various t-tests are compared to the critical values from a standard normal distribution. The rejection rates for the limiting case ($T = \infty$) are based on the asymptotic distributions given in Theorem 5.

Table 3
Survival rates of risk factors: two useful, one unpriced and one useless factors (modified HJ-distance case)

Panel A: Correctly specified model

	Useful ($\gamma_1^* \neq 0$)		Useful ($\gamma_2^* \neq 0$)		Useful ($\gamma_3^* = 0$)		Useless		Prob.	
T	$t_c(\hat{\gamma}_1)$	$t_m(\hat{\gamma}_1)$	$t_c(\hat{\gamma}_2)$	$t_m(\hat{\gamma}_2)$	$t_c(\hat{\gamma}_3)$	$t_m(\hat{\gamma}_3)$	$t_c(\hat{\gamma}_4)$	$t_m(\hat{\gamma}_4)$	MS_c	MS_m
200	0.253	0.239	0.380	0.355	0.010	0.008	0.013	0.001	0.023	0.008
600	0.862	0.852	0.962	0.958	0.010	0.009	0.008	0.000	0.018	0.009
1000	0.986	0.984	0.999	0.999	0.010	0.009	0.006	0.000	0.016	0.009

Panel B: Misspecified model

	Useful ($\gamma_1^* \neq 0$)		Useful ($\gamma_2^* \neq 0$)		Useful ($\gamma_3^* = 0$)		Useless		Prob.	
T	$t_c(\hat{\gamma}_1)$	$t_m(\hat{\gamma}_1)$	$t_c(\hat{\gamma}_2)$	$t_m(\hat{\gamma}_2)$	$t_c(\hat{\gamma}_3)$	$t_m(\hat{\gamma}_3)$	$t_c(\hat{\gamma}_4)$	$t_m(\hat{\gamma}_4)$	MS_c	MS_m
200	0.242	0.213	0.368	0.320	0.013	0.007	0.084	0.005	0.096	0.012
600	0.818	0.776	0.930	0.908	0.013	0.007	0.201	0.006	0.211	0.013
1000	0.958	0.934	0.989	0.983	0.013	0.007	0.295	0.008	0.304	0.015

The table presents the survival rates of the useful and useless factors in a model with a constant, two useful factors (with $\gamma_1^* \neq 0$ and $\gamma_2^* \neq 0$), a useful factor that does not contribute to pricing (with $\gamma_3^* = 0$) and a useless factor (with γ_4^* 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 false discovery rate of the multiple testing procedure is controlled using the Bonferroni method. The last two columns of the table report the probability that at least one useless or unpriced useful factor survives using the t-tests under correctly specified models (MS_c) and misspecification-robust t-tests (MS_m). 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 excess returns on the 25 Fama-French size and book-to-market portfolios and the 17 Fama-French industry portfolios for the period 1959:2–2012:12.

Table 4
Survival rates of risk factors: two useful and two useless factors (modified HJ-distance case)

Panel A: Correctly specified model

	Useful ($\gamma_1^* \neq 0$)		Useful ($\gamma_2^* \neq 0$)		Useless		Useless		Prob.	
T	$t_c(\hat{\gamma}_1)$	$t_m(\hat{\gamma}_1)$	$t_c(\hat{\gamma}_2)$	$t_m(\hat{\gamma}_2)$	$t_c(\hat{\gamma}_3)$	$t_m(\hat{\gamma}_3)$	$t_c(\hat{\gamma}_4)$	$t_m(\hat{\gamma}_4)$	MS_c	MS_m
200	0.272	0.254	0.375	0.344	0.012	0.001	0.012	0.001	0.024	0.001
600	0.891	0.877	0.959	0.951	0.007	0.000	0.007	0.000	0.014	0.001
1000	0.991	0.989	0.999	0.998	0.006	0.000	0.006	0.000	0.012	0.000

Panel B: Misspecified model

	Useful ($\gamma_1^* \neq 0$)		Useful ($\gamma_2^* \neq 0$)		Useless		Useless		Prob.	
T	$t_c(\hat{\gamma}_1)$	$t_m(\hat{\gamma}_1)$	$t_c(\hat{\gamma}_2)$	$t_m(\hat{\gamma}_2)$	$t_c(\hat{\gamma}_3)$	$t_m(\hat{\gamma}_3)$	$t_c(\hat{\gamma}_4)$	$t_m(\hat{\gamma}_4)$	MS_c	MS_m
200	0.252	0.218	0.352	0.294	0.075	0.004	0.075	0.004	0.147	0.008
600	0.812	0.751	0.900	0.857	0.178	0.005	0.179	0.005	0.340	0.010
1000	0.947	0.908	0.976	0.957	0.263	0.006	0.261	0.006	0.482	0.013

The table presents the survival rates of the useful and useless factors in a model with a constant, two useful factors (with $\gamma_1^* \neq 0$ and $\gamma_2^* \neq 0$), and two useless factors (with γ_3^* and γ_4^* 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 false discovery rate of the multiple testing procedure is controlled using the Bonferroni method. The last two columns of the table report the probability that at least one useless factor survives using the t-tests under correctly specified models (MS_c) and misspecification-robust t-tests (MS_m). 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 excess returns on the 25 Fama-French size and book-to-market portfolios and the 17 Fama-French industry portfolios for the period 1959:2–2012:12.

Table 5
Survival rates when a linear combination of the factors is useless (modified HJ-distance case)

Panel A: Correctly specified model

	Both factors survive		One factor survives		No factor survives	
T	t_c	t_m	t_c	t_m	t_c	t_m
200	0.026	0.003	0.247	0.250	0.727	0.747
600	0.015	0.001	0.677	0.685	0.308	0.313
1000	0.013	0.001	0.889	0.900	0.097	0.099

Panel B: Misspecified model

	Both factors survive		One factor survives		No factor survives	
T	t_c	t_m	t_c	t_m	t_c	t_m
200	0.140	0.013	0.228	0.255	0.631	0.733
600	0.275	0.015	0.505	0.684	0.219	0.301
1000	0.377	0.016	0.563	0.890	0.060	0.094

The table presents the probability that both factors survive, only one factor survives, and no factor survives in a model in which a linear combination of two useful factors is useless. The sequential procedure is implemented by using the misspecification-robust t-test (t_m column) as well as the t-test under correctly specified models (t_c column). The false discovery rate of the multiple testing procedure is controlled using the Bonferroni method. 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 excess returns on the 25 Fama-French size and book-to-market portfolios and the 17 Fama-French industry portfolios for the period 1959:2–2012:12.

Table 6
Empirical size of the *t*-tests in a model with a useful factor (optimal GMM case)

Panel A: Correctly specified model

		$\hat{\gamma}_0$			$\hat{\gamma}_1$		
t-test	T	10%	5%	1%	10%	5%	1%
t_c	200	0.176	0.142	0.108	0.114	0.061	0.015
	600	0.140	0.100	0.063	0.103	0.052	0.011
	1000	0.125	0.082	0.043	0.102	0.051	0.010
	∞	0.100	0.050	0.010	0.100	0.050	0.010
t_m	200	0.173	0.141	0.108	0.107	0.055	0.012
	600	0.139	0.100	0.063	0.102	0.051	0.010
	1000	0.125	0.081	0.043	0.101	0.050	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	0.182	0.147	0.110	0.122	0.067	0.018
	600	0.143	0.103	0.065	0.110	0.057	0.013
	1000	0.128	0.085	0.044	0.107	0.055	0.012
	∞	0.100	0.050	0.010	0.100	0.050	0.010
t_m	200	0.175	0.144	0.110	0.109	0.056	0.012
	600	0.140	0.101	0.064	0.103	0.052	0.011
	1000	0.125	0.083	0.044	0.101	0.051	0.010
	∞	0.100	0.050	0.010	0.100	0.050	0.010

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 estimated by optimal (3-step) GMM. 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 size and book-to-market Fama-French portfolio returns, the 17 Fama-French industry portfolio returns and the one-month T-bill rate for the period 1959:2–2012:12. The various t-statistics are compared to the critical values from a standard normal distribution.

Table 7
Empirical size of the t-tests in a model with a useless factor (optimal GMM case)

Panel A: Correctly specified model

		$\hat{\gamma}_0$			$\hat{\gamma}_1$		
t-test	T	10%	5%	1%	10%	5%	1%
t_c	200	0.012	0.004	0.000	0.150	0.088	0.026
	600	0.003	0.000	0.000	0.107	0.053	0.009
	1000	0.002	0.000	0.000	0.100	0.047	0.007
	∞	0.001	0.000	0.000	0.088	0.039	0.005
t_m	200	0.002	0.000	0.000	0.038	0.015	0.002
	600	0.000	0.000	0.000	0.024	0.007	0.000
	1000	0.000	0.000	0.000	0.018	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	0.043	0.020	0.004	0.350	0.267	0.146
	600	0.035	0.013	0.002	0.475	0.391	0.248
	1000	0.040	0.015	0.002	0.559	0.481	0.336
	∞	0.088	0.039	0.005	1.000	1.000	1.000
t_m	200	0.007	0.002	0.000	0.079	0.039	0.009
	600	0.003	0.001	0.000	0.083	0.040	0.007
	1000	0.003	0.000	0.000	0.088	0.043	0.008
	∞	0.001	0.000	0.000	0.100	0.050	0.010

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 estimated by optimal (3-step) GMM. 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 size and book-to-market Fama-French portfolio returns, the 17 Fama-French industry portfolio returns and the one-month T-bill rate for the period 1959:2–2012:12. 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 equivalent to those based on the asymptotic distributions given in Theorem 2.

Table 8
Empirical size of the t-tests in a model with a useful and a useless factor (optimal GMM case)

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	0.064	0.029	0.008	0.118	0.064	0.015	0.153	0.091	0.028
	600	0.061	0.029	0.008	0.101	0.051	0.010	0.108	0.054	0.009
	1000	0.058	0.025	0.006	0.097	0.049	0.009	0.099	0.048	0.007
	∞	0.052	0.020	0.002	0.096	0.047	0.009	0.088	0.039	0.005
t_m	200	0.031	0.013	0.004	0.103	0.052	0.011	0.040	0.016	0.002
	600	0.038	0.017	0.006	0.095	0.047	0.009	0.024	0.006	0.000
	1000	0.037	0.016	0.004	0.092	0.045	0.008	0.021	0.006	0.000
	∞	0.037	0.014	0.002	0.092	0.045	0.008	0.018	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	0.086	0.041	0.010	0.144	0.084	0.026	0.350	0.266	0.144
	600	0.077	0.034	0.007	0.124	0.067	0.016	0.471	0.385	0.241
	1000	0.076	0.032	0.006	0.120	0.065	0.015	0.552	0.473	0.330
	∞	0.088	0.039	0.005	0.088	0.039	0.005	1.000	1.000	1.000
t_m	200	0.026	0.010	0.003	0.106	0.056	0.012	0.081	0.040	0.008
	600	0.018	0.006	0.002	0.089	0.042	0.008	0.082	0.040	0.008
	1000	0.013	0.005	0.001	0.080	0.037	0.006	0.089	0.042	0.008
	∞	0.001	0.000	0.000	0.001	0.000	0.000	0.100	0.050	0.010

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 estimated by optimal (3-step) GMM. 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 size and book-to-market Fama-French portfolio returns, the 17 Fama-French industry portfolio returns and the one-month T-bill rate for the period 1959:2–2012:12. 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 equivalent to those based on the asymptotic distributions given in Theorem 2.

Table 9
Survival rates of risk factors: two useful, one unpriced and one useless factors (optimal GMM case)

Panel A: Correctly specified model

	Useful ($\gamma_1^* \neq 0$)		Useful ($\gamma_2^* \neq 0$)		Useful ($\gamma_3^* = 0$)		Useless		Prob.	
T	$t_c(\hat{\gamma}_1)$	$t_m(\hat{\gamma}_1)$	$t_c(\hat{\gamma}_2)$	$t_m(\hat{\gamma}_2)$	$t_c(\hat{\gamma}_3)$	$t_m(\hat{\gamma}_3)$	$t_c(\hat{\gamma}_4)$	$t_m(\hat{\gamma}_4)$	MS_c	MS_m
200	0.744	0.708	0.812	0.770	0.042	0.030	0.048	0.004	0.087	0.034
600	0.999	0.999	1.000	1.000	0.016	0.014	0.014	0.001	0.029	0.014
1000	1.000	1.000	1.000	1.000	0.015	0.014	0.011	0.000	0.026	0.014

Panel B: Misspecified model

	Useful ($\gamma_1^* \neq 0$)		Useful ($\gamma_2^* \neq 0$)		Useful ($\gamma_3^* = 0$)		Useless		Prob.	
T	$t_c(\hat{\gamma}_1)$	$t_m(\hat{\gamma}_1)$	$t_c(\hat{\gamma}_2)$	$t_m(\hat{\gamma}_2)$	$t_c(\hat{\gamma}_3)$	$t_m(\hat{\gamma}_3)$	$t_c(\hat{\gamma}_4)$	$t_m(\hat{\gamma}_4)$	MS_c	MS_m
200	0.713	0.639	0.785	0.697	0.062	0.033	0.157	0.013	0.207	0.045
600	0.994	0.995	0.997	0.998	0.026	0.014	0.219	0.009	0.237	0.023
1000	0.999	0.999	1.000	1.000	0.023	0.013	0.299	0.009	0.314	0.022

The table presents the survival rates of the useful and useless factors in a model with a constant, two useful factors (with $\gamma_1^* \neq 0$ and $\gamma_2^* \neq 0$), a useful factor that does not contribute to pricing (with $\gamma_3^* = 0$) and a useless factor (with γ_4^* unidentified) estimated by optimal (3-step) GMM. 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 false discovery rate of the multiple testing procedure is controlled using the Bonferroni method. The last two columns of the table report the probability that at least one useless or unpriced useful factor survives using the t-tests under correctly specified models (MS_c) and misspecification-robust t-tests (MS_m). 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 size and book-to-market Fama-French portfolio returns, the 17 Fama-French industry portfolio returns and the one-month T-bill rate for the period 1959:2–2012:12.

Table 10
Survival rates of risk factors: two useful and two useless factors (optimal GMM case)

Panel A: Correctly specified model

	Useful ($\gamma_1^* \neq 0$)		Useful ($\gamma_2^* \neq 0$)		Useless		Useless		Prob.	
T	$t_c(\hat{\gamma}_1)$	$t_m(\hat{\gamma}_1)$	$t_c(\hat{\gamma}_2)$	$t_m(\hat{\gamma}_2)$	$t_c(\hat{\gamma}_3)$	$t_m(\hat{\gamma}_3)$	$t_c(\hat{\gamma}_4)$	$t_m(\hat{\gamma}_4)$	MS_c	MS_m
200	0.749	0.715	0.807	0.768	0.048	0.005	0.047	0.005	0.092	0.009
600	0.999	0.999	1.000	1.000	0.014	0.001	0.014	0.001	0.028	0.001
1000	1.000	1.000	1.000	1.000	0.011	0.000	0.010	0.000	0.021	0.001

Panel B: Misspecified model

	Useful ($\gamma_1^* \neq 0$)		Useful ($\gamma_2^* \neq 0$)		Useless		Useless		Prob.	
T	$t_c(\hat{\gamma}_1)$	$t_m(\hat{\gamma}_1)$	$t_c(\hat{\gamma}_2)$	$t_m(\hat{\gamma}_2)$	$t_c(\hat{\gamma}_3)$	$t_m(\hat{\gamma}_3)$	$t_c(\hat{\gamma}_4)$	$t_m(\hat{\gamma}_4)$	MS_c	MS_m
200	0.356	0.283	0.453	0.359	0.153	0.012	0.152	0.011	0.284	0.023
600	0.843	0.873	0.915	0.935	0.224	0.011	0.222	0.011	0.415	0.022
1000	0.951	0.973	0.977	0.987	0.295	0.012	0.292	0.011	0.533	0.023

The table presents the survival rates of the useful and useless factors in a model with a constant, two useful factors (with $\gamma_1^* \neq 0$ and $\gamma_2^* \neq 0$), and two useless factors (with γ_3^* and γ_4^* unidentified) estimated by optimal (3-step) GMM. 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 false discovery rate of the multiple testing procedure is controlled using the Bonferroni method. The last two columns of the table report the probability that at least one useless factor survives using the t-tests under correctly specified models (MS_c) and misspecification-robust t-tests (MS_m). 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 size and book-to-market Fama-French portfolio returns, the 17 Fama-French industry portfolio returns and the one-month T-bill rate for the period 1959:2–2012:12.

Table 11
Survival rates when a linear combination of the factors is useless (optimal GMM case)

Panel A: Correctly specified model

	Both factors survive		One factor survives		No factor survives	
T	t_c	t_m	t_c	t_m	t_c	t_m
200	0.046	0.006	0.259	0.253	0.695	0.741
600	0.020	0.002	0.673	0.683	0.306	0.315
1000	0.016	0.001	0.887	0.899	0.097	0.099

Panel B: Misspecified model

	Both factors survive		One factor survives		No factor survives	
T	t_c	t_m	t_c	t_m	t_c	t_m
200	0.186	0.017	0.228	0.240	0.586	0.743
600	0.295	0.017	0.489	0.670	0.216	0.313
1000	0.389	0.019	0.552	0.882	0.059	0.099

The table presents the probability that both factors survive, only one factor survives, and no factor survives in a model estimated by optimal (3-step) GMM in which a linear combination of two useful factors is useless. The sequential procedure is implemented by using the misspecification-robust t-test (t_m column) as well as the t-test under correctly specified models (t_c column). The false discovery rate of the multiple testing procedure is controlled using the Bonferroni method. 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 size and book-to-market Fama-French portfolio returns, the 17 Fama-French industry portfolio returns and the one-month T-bill rate for the period 1959:2–2012:12.

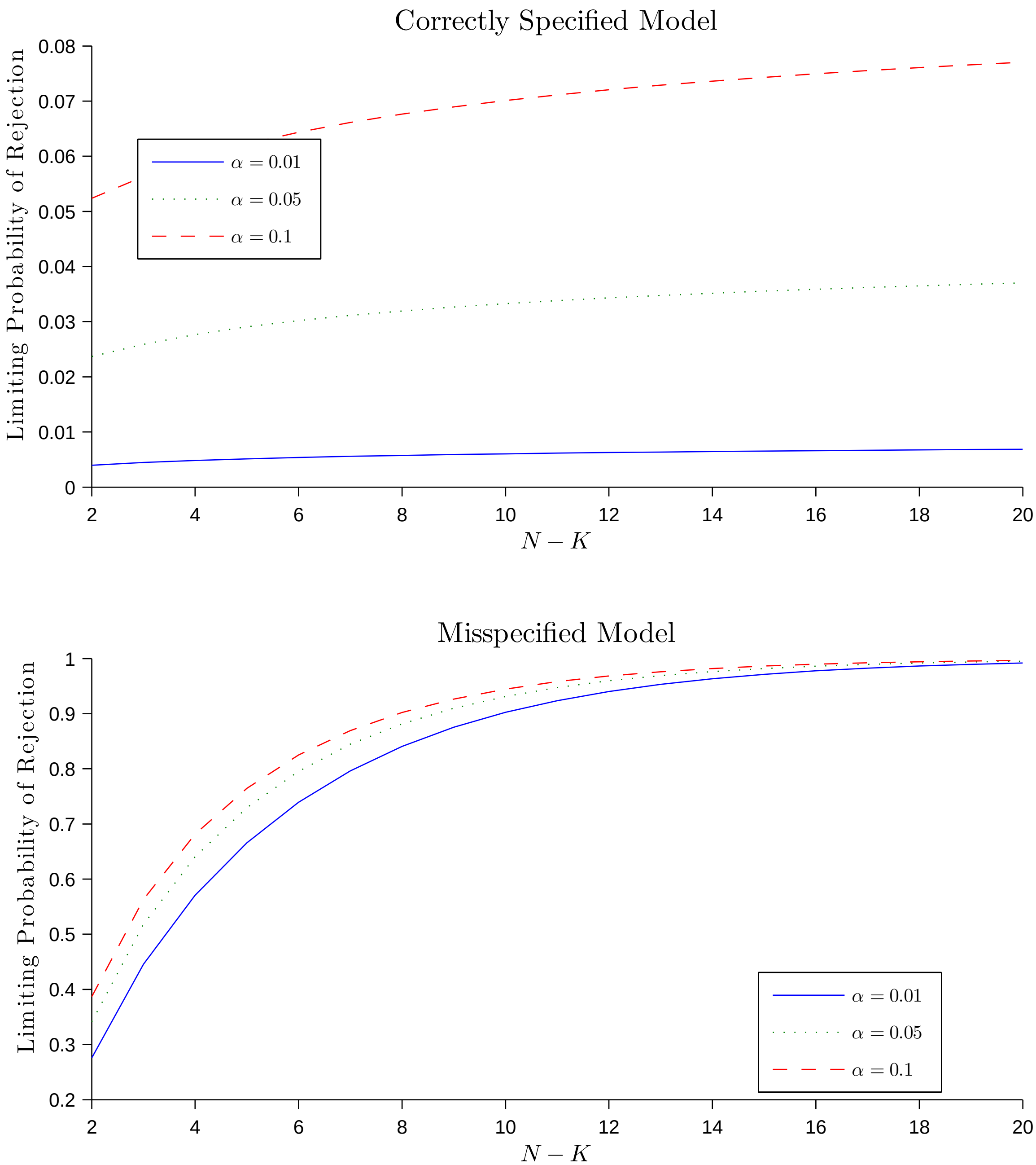

Figure 1

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.