On the estimation of asset pricing models using univariate betas

Raymond Kan and Cesare Robotti*

Abstract

We derive asymptotic standard errors of risk premia estimates based on the popular two-pass cross-sectional regression methodology developed by Black, Jensen, and Scholes (1972) and Fama and MacBeth (1973) when univariate betas are used as regressors. Our standard errors are robust to model misspecification and allow for general distributional assumptions. In testing whether the beta risk of a given factor is priced, our misspecification robust standard error can lead to economically different conclusions from those based on the Jagannathan and Wang (1998) standard error which is derived under the correctly specified model.

Keywords: Asset pricing models; Risk premia; Univariate betas; Model misspecification

JEL classification: G12

*Kan is from the University of Toronto. Robotti is from the Federal Reserve Bank of Atlanta. We thank Esther Eiling and an anonymous referee for helpful discussions and comments and Martin Lettau and Kenneth French for making their data available to us. Kan gratefully acknowledges financial support from the National Bank Financial of Canada, the Social Sciences and Humanities Research Council of Canada, and the Center for Financial Innovation and Stability at the Federal Reserve Bank of Atlanta. The views expressed here are the authors' and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System. Corresponding author: Cesare Robotti, Research Department, Federal Reserve Bank of Atlanta, 1000 Peachtree St. N.E., Atlanta, Georgia, 30309, USA; Tel: (404) 498-8543; Fax: (404) 498-8810; E-mail: cesare.robotti@atl.frb.org.

1. Introduction

The popular two-pass cross-sectional regression (CSR) methodology developed by Black, Jensen, and Scholes (1972) and Fama and MacBeth (1973) is often used for estimating risk premia and testing pricing models that relate expected security returns to asset betas on economic factors (beta pricing models). Usually, asset betas are defined as the ordinary least squares (OLS) slope coefficients in the multiple regression of asset returns on factors and are referred to as *multivariate* or *multiple regression* betas. Unless the factors are uncorrelated, the beta of an asset with respect to a particular factor in general depends on what other factors are included in the first-pass time-series OLS regression. As a result, a factor can possess additional explanatory power for the cross-sectional differences in expected returns but yet have a zero risk premium in a model with multiple factors. This makes it problematic to use the risk premium of a factor for the purpose of model selection. To overcome this problem, Chen, Roll, and Ross (1986) and Jagannathan and Wang (1996, 1998) define the beta of an asset with respect to a given factor as the OLS slope coefficient in a simple regression of its return on the factor. These betas are usually referred to as *univariate* or *simple regression* betas. In models with univariate betas, adding or deleting a factor will not change the values of the betas corresponding to the other factors and selecting models based on risk premia becomes more meaningful.

Jagannathan and Wang (1998) present an asymptotic theory for models with univariate betas. However, their results rest on the assumption that the model is correctly specified. It is difficult to justify this assumption when estimating the zero-beta rate and risk premia parameters from many different models because some (if not all) of the models are bound to be misspecified. The main contribution of this paper is to propose misspecification robust asymptotic standard errors of the estimated zero-beta rate and risk premia in models with univariate betas. In addition, under a multivariate elliptical assumption, we provide simple expressions for the asymptotic variances of the zero-beta rate and risk premia estimates. In the case of the generalized least squares (GLS) CSR estimators, we show that the asymptotic variances are always larger when the model is misspecified.

The paper is organized as follows. Section 2 presents an asymptotic analysis of the zero-beta rate and risk premia estimates for models with univariate betas under potential model misspecification. Section 3 provides an empirical example and Section 4 concludes.

2. Asymptotic analysis under potentially misspecified models

Let f be a K-vector of factors and R a vector of returns on N test assets. We define $Y = [f',\ R']'$ with mean and covariance matrix

$$\mu = E[Y] \equiv \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \tag{1}$$

$$V = \mathrm{Var}[Y] \equiv \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix}, \tag{2}$$

where V is assumed to be positive definite. The univariate betas of the N portfolios w.r.t. the K factors are defined as $\beta^* = V_{21}D^{-1}$, where $D = \mathrm{Diag}(V_{11})$ is a diagonal matrix of the diagonal elements of V_{11}. When the proposed beta pricing model is correctly specified, μ_2 is exactly linear in β^*. As a result, the pricing errors, e, of the N test assets are

$$e \equiv \mu_2 - X^*\gamma^* = 0_N, \tag{3}$$

where $X^* = [1_N,\ \beta^*]$ is assumed to be of full column rank, 0_N is an N-vector of zeros, 1_N is an N-vector of ones, and $\gamma^* = [\gamma_0^*,\ \gamma_1^{*\prime}]'$ is a vector consisting of the zero-beta rate (γ_0^*) and risk premia (γ_1^*). When the model is misspecified, e will be nonzero regardless of the choice of γ^*. Then, γ^* is chosen to minimize some form of aggregate pricing errors. Denoting by W an $N \times N$ symmetric positive definite matrix,

$$\gamma_W^* \equiv \begin{bmatrix} \gamma_{W,0}^* \\ \gamma_{W,1}^* \end{bmatrix} = \mathrm{argmin}_{\gamma^*}(\mu_2 - X^*\gamma^*)'W(\mu_2 - X^*\gamma^*) = (X^{*\prime}WX^*)^{-1}X^{*\prime}W\mu_2 \tag{4}$$

and

$$e_W = \mu_2 - X^*\gamma_W^*. \tag{5}$$

Unless the model is correctly specified, γ_W^* and e_W depend on the choice of W. Popular choices of W are $W = I_N$ (OLS CSR), $W = V_{22}^{-1}$ (GLS CSR), and $W = \Sigma_d^{-1}$ (weighted least squares (WLS) CSR), where $\Sigma_d = \mathrm{Diag}(\Sigma)$ and $\Sigma = V_{22} - V_{21}V_{11}^{-1}V_{12}$. To simplify the notation, we suppress the subscript W from γ_W^* and e_W when the choice of W is clear from the context.

Let $Y_t = [f_t',\ R_t']'$, where f_t and R_t are the vectors of factors and returns at time t, respectively. Assume that Y_t is jointly stationary and ergodic with finite fourth moment. Suppose we have T

observations on Y_t and denote the sample moments of Y_t by

$$\hat{\mu} \equiv \begin{bmatrix} \hat{\mu}_1 \\ \hat{\mu}_2 \end{bmatrix} = \frac{1}{T}\sum_{t=1}^{T} Y_t, \tag{6}$$

$$\hat{V} \equiv \begin{bmatrix} \hat{V}_{11} & \hat{V}_{12} \\ \hat{V}_{21} & \hat{V}_{22} \end{bmatrix} = \frac{1}{T}\sum_{t=1}^{T}(Y_t - \hat{\mu})(Y_t - \hat{\mu})'. \tag{7}$$

Letting $\hat{\beta}^* = \hat{V}_{21}\hat{D}^{-1}$, when the weighting matrix W is known, we can estimate γ_W^* in (4) by

$$\hat{\gamma}^* = (\hat{X}^{*\prime} W \hat{X}^*)^{-1}\hat{X}^{*\prime} W \hat{\mu}_2, \tag{8}$$

where $\hat{X}^* = [1_N,\ \hat{\beta}^*]$. In the GLS and WLS cases, we need to substitute a consistent estimate of W, say $\hat{W}$, in (5) and (8) (e.g., $\hat{W} = \hat{V}_{22}^{-1}$ for GLS and $\hat{W} = \hat{\Sigma}_d^{-1}$ for WLS). The following proposition presents the asymptotic distribution of $\hat{\gamma}^*$.

Proposition 1. *Under a potentially misspecified model, the asymptotic distribution of $\hat{\gamma}^*$ is given by*

$$\sqrt{T}(\hat{\gamma}^* - \gamma^*) \overset{A}{\sim} N(0_{K+1}, V(\hat{\gamma}^*)), \tag{9}$$

where

$$V(\hat{\gamma}^*) = \sum_{j=-\infty}^{\infty} E[h_t h_{t+j}']. \tag{10}$$

To simplify the h_t expressions in the different CSRs, we define $H^* = (X^{*\prime} W X^*)^{-1}$, $A^* = H^* X^{*\prime} W$, $\gamma_t^* = A^* R_t$, $z_t^* = [0,\ (f_t - \mu_1)' D^{-1}]'$, $D_t = \mathrm{Diag}((f_t - \mu_1)(f_t - \mu_1)')$, $G_t^* = [\beta^* D_t - (R_t - \mu_2)(f_t - \mu_1)']$, and $u_t = e' W (R_t - \mu_2)$, where W equals V_{22}^{-1} in the GLS case and Σ_d^{-1} in the WLS case.

(1) In the known weighting matrix W case,

$$h_t = (\gamma_t^* - \gamma^*) + A^* G_t^* D^{-1} \gamma_1^* + H^* z_t^* u_t. \tag{11}$$

(2) In the GLS case,

$$h_t = (\gamma_t^* - \gamma^*) + A^* G_t^* D^{-1} \gamma_1^* + H^* z_t^* u_t - (\gamma_t^* - \gamma^*) u_t. \tag{12}$$

(3) In the WLS case,

$$h_t = (\gamma_t^* - \gamma^*) + A^* G_t^* D^{-1} \gamma_1^* + H^* z_t^* u_t - A^* \Psi_t \Sigma_d^{-1} e, \tag{13}$$

where $\Psi_t = \text{Diag}(\epsilon_t \epsilon_t')$, *and* $\epsilon_t = R_t - \mu_2 - V_{21}V_{11}^{-1}(f_t - \mu_1)$.

When the model is correctly specified, we have

$$h_t = (\gamma_t^* - \gamma^*) + A^* G_t^* D^{-1} \gamma_1^*. \tag{14}$$

Proof. See Appendix.

In the known W case, we can easily construct a consistent estimator of $V(\hat{\gamma}^*)$ by replacing h_t with

$$\hat{h}_t = (\hat{\gamma}_t^* - \hat{\gamma}^*) + \hat{A}^* \hat{G}_t^* \hat{D}^{-1} \hat{\gamma}_1^* + \hat{H}^* \hat{z}_t^* \hat{u}_t, \tag{15}$$

where $\hat{H}^* = (\hat{X}^{*\prime} W \hat{X}^*)^{-1}$, $\hat{A}^* = \hat{H}^* \hat{X}^{*\prime} W$, $\hat{\gamma}_t^* = \hat{A}^* R_t$, $\hat{u}_t = \hat{e}' W (R_t - \hat{\mu}_2)$, $\hat{D} = \text{Diag}(\hat{V}_{11})$, $\hat{D}_t = \text{Diag}[(f_t - \hat{\mu}_1)(f_t - \hat{\mu}_1)']$, $\hat{G}_t^* = [\hat{\beta}^* \hat{D}_t - (R_t - \hat{\mu}_2)(f_t - \hat{\mu}_1)']$, and $\hat{z}_t^* = [0,\ (f_t - \hat{\mu}_1)' \hat{D}^{-1}]'$. Similarly, one needs to replace the population quantities in (12)–(13) with their sample counterparts to obtain a consistent estimator of $V(\hat{\gamma}^*)$ in the GLS and WLS cases.[1] Note that for the case of correctly specified model, our expressions in (10) and (14) provide a substantial simplification of the corresponding expressions in Theorem 7 of Jagannathan and Wang (1998).

An inspection of (11) reveals that there are three sources that contribute to the asymptotic variance of $\hat{\gamma}^*$. The first term, $\gamma_t^* - \gamma^*$, measures the asymptotic variance of $\hat{\gamma}^*$ when the *true* betas are used in the CSR. For example, when R_t is i.i.d., then γ_t^* is also i.i.d. and we can use the time-series variance of γ_t^* to compute the standard error of $\hat{\gamma}^*$ (see Fama and MacBeth (1973)). However, since the estimated β^* is used in the second-pass CSR, there is an errors-in-variables (EIV) problem and the second term, $A^* G_t^* D^{-1} \gamma_1^*$, represents the necessary EIV adjustment. These two terms together give us $V(\hat{\gamma}^*)$ under the correctly specified model. When the model is misspecified, there is a third term $H^* z_t^* u_t$, the misspecification adjustment term, that contributes to the asymptotic variance of $\hat{\gamma}^*$. This term has been ignored by Jagannathan and Wang (1998) and other researchers.[2] To better understand the importance of the misspecification adjustment term, in the following lemma we derive an explicit expression for $V(\hat{\gamma}^*)$.

Lemma 1. *Suppose that factors and returns are i.i.d. multivariate elliptically distributed with kur-*

[1] If h_t is uncorrelated over time, then $V(\hat{\gamma}^*) = E[h_t h_t']$ and $\hat{V}(\hat{\gamma}^*) = \frac{1}{T}\sum_{t=1}^{T} \hat{h}_t \hat{h}_t'$. When h_t is autocorrelated, one can use the method of Newey and West (1987).

[2] In the estimated GLS and WLS cases, the misspecification adjustment term contains the additional quantities $-(\gamma_t^* - \gamma^*)u_t$ and $-A^* \Psi_t \Sigma_d^{-1} e$, respectively. These additional terms are due to the estimation error in $\hat{W}$.

tosis parameter κ.[3] *Let* $\tilde{D} = \begin{bmatrix} 0 & 0_K' \\ 0_K & D^{-1}V_{11}D^{-1} \end{bmatrix}$ *and* $\odot$ *denote the Hadamard product. Define*

$$\begin{aligned} \Upsilon_w &= [1+(1+\kappa)(\gamma_1^{*\prime}D^{-1}V_{11}D^{-1}\gamma_1^*)]A^*V_{22}A^{*\prime} + (1+\kappa)\times \\ & \begin{bmatrix} 0 & 0_K' \\ 0_K & 2(D^{-1}\gamma_1^*\gamma_1^{*\prime}D^{-1})\odot V_{11}\odot V_{11} - 4\mathrm{Diag}(\gamma_1^*\gamma_1^{*\prime}D^{-1}V_{11}) + \gamma_1^*\gamma_1^{*\prime} \end{bmatrix}, \quad (16) \\ \Upsilon_{w1} &= -(1+\kappa)A^*V_{22}We[0,\ \gamma_1^{*\prime}D^{-1}V_{11}D^{-1}]H^*, \quad (17) \end{aligned}$$

where $W = V_{22}^{-1}$ *in the GLS case and* $W = \Sigma_d^{-1}$ *in the WLS case. The asymptotic variance of* $\hat{\gamma}^*$ *is given by*

$$V(\hat{\gamma}^*) = \Upsilon_w + \Upsilon_{w1} + \Upsilon_{w1}' + \Upsilon_{w2}, \quad (18)$$

where Υ_w *is the asymptotic variance of* $\hat{\gamma}^*$ *when the model is correctly specified, and* $\Upsilon_{w1}+\Upsilon_{w1}'+\Upsilon_{w2}$ *is the adjustment term due to model misspecification.*

(1) In the known weighting matrix W *case,*

$$\Upsilon_{w2} = (1+\kappa)(e'WV_{22}We)H^*\tilde{D}H^*. \quad (19)$$

(2) In the GLS case, Υ_{w1} *vanishes and*

$$\Upsilon_{w2} = (1+\kappa)(e'V_{22}^{-1}e)(\tilde{H}\tilde{D}\tilde{H} + \tilde{H}), \quad (20)$$

where $\tilde{H} = (X^{*\prime}\Sigma^{-1}X^*)^{-1}$.

(3) In the WLS case,

$$\Upsilon_{w2} = (1+\kappa)\left[(e'\Sigma_d^{-1}V_{22}\Sigma_d^{-1}e)H^*\tilde{D}H^* + 2A^*\Phi A^{*\prime}\right], \quad (21)$$

where Φ *is an* $(N\times N)$ *matrix with its* (i,j)*-th element equal to* $\rho_{ij}^2 e_i e_j$ *and* $\rho_{ij} = \mathrm{Corr}[\epsilon_{it}, \epsilon_{jt}]$.

Proof. See Appendix.

In the known W and WLS cases, the misspecification adjustment term is not necessarily positive semidefinite. In contrast, in the GLS case, the misspecification adjustment term is positive definite and can be rewritten as

$$(1+\kappa)(e'V_{22}^{-1}e)H_{22}^*D^{-1}[V_{11} - V_{12}V_{22}^{-1}V_{21} + V_{12}V_{22}^{-1}1_N(1_N'V_{22}^{-1}1_N)^{-1}1_N'V_{22}^{-1}V_{21}]D^{-1}H_{22}^*, \quad (22)$$

[3]The kurtosis parameter for an elliptical distribution is defined as $\kappa = \mu_4/(3\sigma^4) - 1$, where σ^2 and μ_4 are the second and fourth central moments of the elliptical distribution, respectively.

where H^*_{22} is the lower right $K \times K$ submatrix of H^*. Therefore, this adjustment is positively related to $e'V_{22}^{-1}e$ and κ, and depends on the correlation between factors and returns through the term $V_{11} - V_{12}V_{22}^{-1}V_{21}$. For factors that have very low correlation with returns, the impact of this term and hence of the misspecification adjustment on the asymptotic variance of $\hat{\gamma}^*_1$ can be large.

3. An empirical example

We apply our methodology to the same data and asset pricing models considered by Lettau and Ludvigson (2001).[4] The results for the Fama and French (1993) three-factor model (FF3) indicate that the Fama and MacBeth (1973), the Jagannathan and Wang (1998), and our misspecification robust t-ratios based on Proposition 1 are all close as the factors are mimicked well by the returns on the test assets. For example, we find that the OLS t-ratios associated with the market factor in FF3 are, in the order, 0.77, 0.73, and 0.57. However, when we consider models with scaled factors such as the scaled capital asset pricing model proposed by Lettau and Ludvigson (2001), the picture changes substantially. For the scaled market factor (i.e., the market factor scaled by the lagged consumption-wealth ratio (CAY)), the OLS t-ratios of Fama and MacBeth (1973) and Jagannathan and Wang (1998) are 3.63 and 2.70, respectively. But, once we account for potential model misspecification, the t-ratio goes down to 1.27. The GLS and WLS results deliver a similar message. In summary, ignoring potential model misspecification can lead to the incorrect conclusion that a given risk factor is priced.

4. Conclusion

We propose a simple methodology for computing misspecification robust asymptotic standard errors of risk premia estimates in models with univariate betas. A nice feature of the proposed standard errors is that they can be used whether the model is correctly specified or not. We show empirically that some factors commonly used in the literature are no longer priced once model misspecification is taken into account.

[4] Detailed estimation results are available from the authors upon request.

Appendix

Proof of Proposition 1: In the following, we provide the proof of Proposition 1 in the estimated GLS and WLS cases as the proof in the known weighting matrix W case is very similar. The proof relies on the fact that $\hat{\gamma}^*$ is a smooth function of $\hat{\mu}$ and $\hat{V}$. Therefore, once we have the asymptotic distribution of $\hat{\mu}$ and $\hat{V}$, we can use the delta method to obtain the asymptotic distribution of $\hat{\gamma}^*$. Let

$$\varphi = \begin{bmatrix} \mu \\ \mathrm{vec}(V) \end{bmatrix}, \qquad \hat{\varphi} = \begin{bmatrix} \hat{\mu} \\ \mathrm{vec}(\hat{V}) \end{bmatrix}. \tag{A1}$$

We first note that $\hat{\mu}$ and $\hat{V}$ can be written as the GMM estimator that uses the moment conditions $E[r_t(\varphi)] = 0_{(N+K)(N+K+1)}$, where

$$r_t(\varphi) = \begin{bmatrix} Y_t - \mu \\ \mathrm{vec}((Y_t - \mu)(Y_t - \mu)' - V) \end{bmatrix}. \tag{A2}$$

Assuming that Y_t is stationary and ergodic with finite fourth moments, we have[5]

$$\sqrt{T}(\hat{\varphi} - \varphi) \overset{A}{\sim} N(0_{(N+K)(N+K+1)}, S_0), \tag{A3}$$

where

$$S_0 = \sum_{j=-\infty}^{\infty} E[r_t(\varphi) r_{t+j}(\varphi)']. \tag{A4}$$

Using the delta method, the asymptotic distribution of $\hat{\gamma}^*$ under the misspecified model is given by

$$\sqrt{T}(\hat{\gamma}^* - \gamma^*) \quad \overset{A}{\sim} \quad N\left(0_{K+1}, \left[\frac{\partial \gamma^*}{\partial \varphi'}\right] S_0 \left[\frac{\partial \gamma^*}{\partial \varphi'}\right]'\right). \tag{A5}$$

In both the GLS and the WLS cases, we have

$$\frac{\partial \gamma^*}{\partial \mu_1'} = 0_{(K+1)\times K}, \qquad \frac{\partial \gamma^*}{\partial \mu_2'} = A^*. \tag{A6}$$

In the GLS case, the derivative of $\gamma^* = H^* X^{*\prime} V_{22}^{-1} \mu_2$ w.r.t. vec(V) is given by

$$\begin{aligned} \frac{\partial \gamma^*}{\partial \mathrm{vec}(V)'} \quad = \quad & \left[H^*[0_K,\ D^{-1}]',\ 0_{(K+1)\times N}\right] \otimes [0_K',\ e'V_{22}^{-1}] - [\gamma_1^{*\prime} D^{-1},\ e'V_{22}^{-1}] \otimes [0_{(K+1)\times K},\ A^*] \\ & + \left(\gamma_1^{*\prime} D^{-1} \otimes A^* \beta^*\right) \Theta_1 \left([I_K,\ 0_{K\times N}] \otimes [I_K,\ 0_{K\times N}]\right), \end{aligned} \tag{A7}$$

[5]Note that S_0 is a singular matrix as $\hat{V}$ is symmetric, so there are redundant elements in $\hat{\varphi}$. We could have written $\hat{\varphi}$ as $[\hat{\mu}',\ \mathrm{vech}(\hat{V})']'$, but the results are the same under both specifications.

where Θ_1 is a $K^2 \times K^2$ matrix such that $\text{vec}(D) = \Theta_1 \text{vec}(V_{11})$.[6] Using the above expression of $\partial\gamma^*/\partial\varphi'$, we can simplify the asymptotic variance of $\hat{\gamma}^*$ to

$$V(\hat{\gamma}^*) = \sum_{j=-\infty}^{\infty} E[h_t(\varphi)h'_{t+j}(\varphi)], \tag{A8}$$

where

$$h_t(\varphi) = \frac{\partial\gamma^*}{\partial\varphi'} r_t(\varphi) = (\gamma_t^* - \gamma^*) + A^* G_t^* D^{-1}\gamma_1^* + H^* z_t^* u_t - (\gamma_t^* - \gamma^*)u_t. \tag{A9}$$

In the WLS case, the derivative of $\gamma^* = H^* X^{*\prime}\Sigma_d^{-1}\mu_2$ w.r.t. $\text{vec}(V)$ is given by

$$\begin{aligned}
\frac{\partial\gamma^*}{\partial\text{vec}(V)'} &= \left[H^*[0_K,\ D^{-1}]',\ 0_{(K+1)\times N}\right] \otimes [0'_K,\ e'\Sigma_d^{-1}] - [\gamma_1^{*\prime} D^{-1},\ 0'_N] \otimes [0_{(K+1)\times K},\ A^*] \\
&\quad + \left(\gamma_1^{*\prime} D^{-1} \otimes A^*\beta^*\right)\Theta_1\left([I_K,\ 0_{K\times N}] \otimes [I_K,\ 0_{K\times N}]\right) \\
&\quad - (e'\Sigma_d^{-1} \otimes A^*)\Theta\left([-\beta,\ I_N] \otimes [-\beta,\ I_N]\right),
\end{aligned} \tag{A10}$$

where Θ is an $N^2 \times N^2$ matrix such that $\text{vec}(\Sigma_d) = \Theta\text{vec}(\Sigma)$. Using the above expression of $\partial\gamma^*/\partial\varphi'$, we can simplify the asymptotic variance of $\hat{\gamma}^*$ to

$$V(\hat{\gamma}^*) = \sum_{j=-\infty}^{\infty} E[h_t(\varphi)h'_{t+j}(\varphi)], \tag{A11}$$

where

$$h_t(\varphi) = \frac{\partial\gamma^*}{\partial\varphi'} r_t(\varphi) = (\gamma_t^* - \gamma^*) + A^* G_t^* D^{-1}\gamma_1^* + H^* z_t^* u_t - A^*\Psi_t\Sigma_d^{-1}e. \tag{A12}$$

Note that when the model is correctly specified, we have $e = 0_N$, so $u_t = 0$ and the $h_t(\varphi)$ in both the GLS and the WLS cases can be simplified to $h_t(\varphi) = (\gamma_t^* - \gamma^*) + A^* G_t^* D^{-1}\gamma_1^*$. This completes the proof.

Proof of Lemma 1: We rely on the mixed moments of multivariate elliptical distributions.[7] Starting from the known weighting matrix case, the asymptotic variance of $\hat{\gamma}^*$ is given by

$$V(\hat{\gamma}^*) = E[h_t h'_t], \tag{A13}$$

where

$$h_t = h_{1t} + h_{2t} + h_{3t}, \tag{A14}$$

[6] Specifically, Θ_1 is a matrix with its (i,i)-th element equals to one, where $i = 1, 1+1(K+1), 1+2(K+1), \ldots, 1+(K-1)(K+1)$, and zero elsewhere.

[7] See, for example, Lemma 2 of Maruyama and Seo (2003).

with

$$h_{1t} = A^*(R_t - \mu_2), \qquad h_{2t} = A^*[\beta^* D_t - (R_t - \mu_2)(f_t - \mu_1)']D^{-1}\gamma_1^*, \qquad h_{3t} = H^* z_t^* u_t. \quad \text{(A15)}$$

It can be shown that the means of h_{1t} to h_{3t} are all equal to zero and

$$E[h_{1t}h_{1t}'] = A^* V_{22} A^{*\prime}. \quad \text{(A16)}$$

In addition, h_{1t} is uncorrelated with h_{2t} and h_{3t}. For h_{2t}, using that $R_t - \mu_2 = \beta(f_t - \mu_1) + \epsilon_t = \beta^* D V_{11}^{-1}(f_t - \mu_1) + \epsilon_t$ and that $A^*\beta^* = [0_K,\ I_K]'$, we have that $h_{2t} = \begin{bmatrix} 0 \\ q_t \end{bmatrix} - A^*\epsilon_t(f_t - \mu_1)' D^{-1}\gamma_1^*$, where $q_t = [D_t - DV_{11}^{-1}(f_t - \mu_1)(f_t - \mu_1)']D^{-1}\gamma_1^*$. Since ϵ_t and f_t are uncorrelated and since u_t is uncorrelated with f_t, we have

$$\begin{aligned} E[h_{2t}h_{2t}'] &= (1+\kappa)(\gamma_1^{*\prime} D^{-1} V_{11} D^{-1}\gamma_1^*) A^* V_{22} A^{*\prime} + (1+\kappa) \times \\ & \begin{bmatrix} 0 & 0_K' \\ 0_K & 2(D^{-1}\gamma_1^*\gamma_1^{*\prime}D^{-1}) \odot V_{11} \odot V_{11} - 4\mathrm{Diag}(\gamma_1^*\gamma_1^{*\prime}D^{-1}V_{11}) + \gamma_1^*\gamma_1^{*\prime} \end{bmatrix}, \quad \text{(A17)} \\ E[h_{3t}h_{3t}'] &= H^* \begin{bmatrix} 0 & 0_K' \\ 0_K & (1+\kappa)(e'WV_{22}We)D^{-1}V_{11}D^{-1} \end{bmatrix} H^*, \quad \text{(A18)} \\ E[h_{2t}h_{3t}'] &= -(1+\kappa)A^* V_{22} We[0,\ \gamma_1^{*\prime} D^{-1} V_{11} D^{-1}]H^*. \quad \text{(A19)} \end{aligned}$$

Collecting terms, we obtain

$$\begin{aligned} \Upsilon_w &= E[h_{1t}h_{1t}'] + E[h_{2t}h_{2t}'] \\ &= [1 + (1+\kappa)(\gamma_1^{*\prime} D^{-1} V_{11} D^{-1}\gamma_1^*)] A^* V_{22} A^{*\prime} + (1+\kappa) \times \\ & \begin{bmatrix} 0 & 0_K' \\ 0_K & 2(D^{-1}\gamma_1^*\gamma_1^{*\prime}D^{-1}) \odot V_{11} \odot V_{11} - 4\mathrm{Diag}(\gamma_1^*\gamma_1^{*\prime}D^{-1}V_{11}) + \gamma_1^*\gamma_1^{*\prime} \end{bmatrix}, \quad \text{(A20)} \\ \Upsilon_{w1} &= E[h_{2t}h_{3t}'] = -(1+\kappa)A^* V_{22} We[0,\ \gamma_1^{*\prime} D^{-1} V_{11} D^{-1}]H^*, \quad \text{(A21)} \\ \Upsilon_{w2} &= E[h_{3t}h_{3t}'] = (1+\kappa)(e'WV_{22}We)H^*\tilde{D}H^*. \quad \text{(A22)} \end{aligned}$$

Turning to the GLS case, we have

$$h_t = h_{1t} + h_{2t} + h_{3t} + h_{4t}, \quad \text{(A23)}$$

where $h_{4t} = -(\gamma_t^* - \gamma^*)u_t$ and h_{1t} to h_{3t} are the same as those in the known weighting matrix case after setting $W = V_{22}^{-1}$. It follows that the Υ_w and Υ_{w1} expressions are the same as the ones in the known weighting matrix case. In the GLS case, Υ_{w1} is a zero matrix because $A^* V_{22} V_{22}^{-1} e = A^* e = 0_{K+1}$. It can be shown that h_{1t} and h_{2t} are uncorrelated with h_{4t}. In addition,

$$E[h_{3t}h_{3t}'] = (1+\kappa)(e'V_{22}^{-1}e)H^*\tilde{D}H^*, \tag{A24}$$

$$E[h_{3t}h_{4t}'] = -(1+\kappa)(e'V_{22}^{-1}e)H^*\begin{bmatrix} 0 & 0_K' \\ 0_K & I_K \end{bmatrix}, \tag{A25}$$

$$E[h_{4t}h_{4t}'] = (1+\kappa)(e'V_{22}^{-1}e)H^*. \tag{A26}$$

Collecting terms and using that $H^* = \tilde{H} + \begin{bmatrix} 0 & 0_K' \\ 0_K & DV_{11}^{-1}D \end{bmatrix}$, we obtain

$$\Upsilon_{w2} = E[h_{3t}h_{3t}'] + E[h_{3t}h_{4t}'] + E[h_{4t}h_{3t}'] + E[h_{4t}h_{4t}'] = (1+\kappa)(e'V_{22}^{-1}e)(\tilde{H}\tilde{D}\tilde{H} + \tilde{H}). \tag{A27}$$

Finally, in the WLS case, we have

$$h_t = h_{1t} + h_{2t} + h_{3t} + h_{4t}, \tag{A28}$$

where $h_{4t} = -A^*\Psi_t\Sigma_d^{-1}e$ and h_{1t} to h_{3t} are the same as those in the known weighting matrix case after setting $W = \Sigma_d^{-1}$. It follows that the Υ_w and Υ_{w1} expressions are the same as the ones in the known weighting matrix case. Since h_{4t} is uncorrelated with h_{1t} to h_{3t}, $E[\Psi_t\Sigma_d^{-1}ee'\Sigma_d^{-1}\Psi_t] = (1+\kappa)(2\Phi + ee')$, and $A^*e = 0_{K+1}$, we obtain

$$E[h_{4t}h_{4t}'] = 2(1+\kappa)A^*\Phi A^{*\prime} \tag{A29}$$

and

$$\Upsilon_{w2} = E[h_{3t}h_{3t}'] + E[h_{4t}h_{4t}'] = (1+\kappa)\left[(e'\Sigma_d^{-1}V_{22}\Sigma_d^{-1}e)H^*\tilde{D}H^* + 2A^*\Phi A^{*\prime}\right]. \tag{A30}$$

When the model is correctly specified, $e = 0_N$ and as a result both Υ_{w1} and Υ_{w2} vanish and we have $V(\hat{\gamma}^*) = \Upsilon_w$. This completes the proof.

References

Black, F., M. Jensen and M. Scholes, 1972, The capital asset pricing model: Some empirical findings, in: M. C. Jensen, ed., Studies in the Theory of Capital Markets (Praeger, New York).

Chen, N., R. Roll and S. Ross, 1986, Economic forces and the stock market, Journal of Business 59, 383–404.

Fama, E. and J. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, Journal of Political Economy 71, 607–636.

Fama, E. and K. French, 1993, Common risk factors in the returns on stocks and bonds, Journal of Financial Economics 33, 3–56.

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

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

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

Maruyama, Y. and T. Seo, 2003, Estimation of moment parameter in elliptical distributions, Journal of the Japan Statistical Society 33, 215–229.

Newey, W. and K. West, 1987, A simple positive definite heteroskedasticity and autocorrelation consistent covariance matrix, Econometrica 55, 703–708.