Factor Model Comparisons with Conditioning Information

Wayne E. Ferson, Andrew F. Siegel, and Junbo L. Wang*

November 30, 2023

Abstract
We develop methods for testing factor models when the weights in portfolios of factors and test assets can vary with lagged information. We derive and evaluate consistent standard errors and finite sample bias adjustments for unconditional maximum squared Sharpe ratios and their differences. Bias adjustment using a second-order approximation performs well. We derive optimal zero beta rates for models with dynamically trading portfolios. Factor models’ Sharpe ratios are larger but standard test asset portfolios’ maximum Sharpe ratios are larger still when there is dynamic trading. As a result, most of the popular factor models are rejected.

*Wayne Ferson is a Professor Emeritus of Finance at University of Southern California Marshall School of Business, and a Research Associate at the National Bureau of Economic Research. He may be reached at 310374-5030, ferson@usc.edu. Andrew F. Siegel is Professor Emeritus at University of Washington Foster School of Business, Box 353200, Seattle, WA. 98195-3200, asiegel@uw.edu. Junbo Wang is an Assistant Professor at Louisiana State University Ourso College of Business, junbowang@lsu.edu. We are grateful to Andrew Detzel, Raymond Kan, Xiaolu Wang, Baozhong Yang, Paolo Zaffaroni, Guofu Zhou, the editor, and an anonymous referee for comments, and to seminar participants at the 2019 Financial Management Conference, the 2021 China International Finance Conference, the 2022 Southwestern Finance Association, the 2022 Southern Finance Association, the 5th World Symposium on Investment Research, the University of Texas at Austin, the University of New Orleans, the Louisiana State University, and the University of Washington in Seattle.
I. Introduction

Classical tests of asset pricing models examine portfolio efficiency by comparing squared Sharpe ratios (e.g., Sharpe, 1988), the ratios of expected excess return to standard deviation. The tests of Gibbons, Ross, and Shanken (GRS, 1989) compare the maximum squared Sharpe ratio, \( S^2(r,f) \), of a portfolio formed from test assets \( r \) and factors \( f \), to that of the portfolio of factors, \( S^2(f) \). The difference between \( S^2(r,f) \) and \( S^2(f) \) is a quadratic form in the factor model’s alphas for the test assets. If the two Sharpe ratios are equal, the test assets have zero alphas in the factor model and the factor portfolio is mean-variance efficient. Dividing the difference of the squared Sharpe ratios by \([1+S^2(f)]\) leads to a test statistic with an exact F-distribution when normality is assumed.


The above papers restrict their analyses in two important dimensions. First, the portfolios of the factors and test assets whose Sharpe ratios are compared must have fixed weights over time. This is not realistic, as most portfolio weights vary over time. Second, these studies measure returns in excess of a short-term Treasury bill rate, pinning the zero-beta rate to the average bill rate. This is not realistic either, and many studies find zero beta rates larger than the average bill rate. This paper removes both of these unrealistic restrictions.
We allow stylized “dynamic trading” in a discrete-time model, using monthly returns and lagged instruments. As the lagged instruments vary over the months the portfolio weights change month-to-month as functions of the lagged information variables. We derive estimates of an optimal zero beta rate for squared Sharpe ratio differences in this setting. We derive asymptotic standard errors, recognizing that the zero beta rate is estimated with sampling error.

We develop a simple approach to testing factor models, with or without dynamic trading. Maximized Sharpe ratios are biased in finite samples, so we provide methods for bias adjustment when there is dynamic trading. Bias-adjusted squared Sharpe ratios or their differences, divided by asymptotic standard errors, form a “t-ratio.” The absolute t-ratio is asymptotically distributed as a $\chi(1)$ variable and simulations show that it may be evaluated using standard rules of thumb; e.g, approximate significance at the 5% level if the t-ratio is larger than two.

We illustrate applications with tests on the Capital Asset Pricing Model (CAPM; Sharpe, 1964), the Fama and French (1996) three-factor model (FF3), the Fama and French (2016) five-factor model (FF5), a six-factor model that appends a momentum factor (FF6), the four-factor investment model of Hou, Xue, and Zhang (Q4, 2015), and a five-factor extension from Hou, Mo, Xue, and Zhang (Q5, 2021) that adds an expected growth factor. In the Internet Appendix, we provide results for mimicking portfolios when factors are not traded, and evidence for the Chen, Roll, and Ross (1986) nontraded factors, consumption growth and a broker-dealer leverage factor from Adrian, Ettula, and Muir (2014).

We find that dynamic trading improves the maximum Sharpe ratios of portfolios of the models’ factors. The Q4 and Q5 model factors are the least affected. The larger impact of dynamic trading is to jack up the maximum Sharpe ratios of the most popular portfolio designs (courtesy of Kenneth French), when treated as the test assets. All of the factor models listed
above are rejected in these portfolios, implying that no portfolio of the model factors is minimum-variance efficient, even with dynamic trading. Direct comparisons reveal that the Q4 model outperforms the FF5 model with or without dynamic trading, and Q5 beats FF6. The evidence for pricing improvement by adding a momentum factor to the FF3 model is strong, but the evidence for adding momentum to the FF5 model appears weak.

When the zero beta rate is estimated using the factor models and the standard test portfolios, its value is larger than the historical average of a US Treasury bill return. The CAPM produces a larger zero beta rate than the other factor models. The zero beta rate can reflect borrowing costs or the premiums of missing factors in the models. The results are consistent with missing factors in the CAPM.

The rest of the paper is organized as follows. Section II presents an overview of Sharpe ratio comparisons. Section III develops our asymptotic results and bias correction for squared Sharpe ratios. The data are described in Section IV. Section V presents simulation results. Section VI presents empirical tests and comparisons of factor pricing models. Section VII concludes the paper. An appendix provides additional material, and an Internet Appendix presents the longer proofs and more.

II. Testing Asset Pricing Models with Squared Sharpe Ratios

Asset pricing models identify a portfolio that is minimum-variance efficient1 In a multiple-factor model, this is a combination of the model’s K traded factors, f, or their

1 The Capital Asset Pricing Model (CAPM, Sharpe (1964)) implies that the market portfolio should be mean-variance efficient. Multiple-beta asset pricing models such as Merton (1973) imply that a combination of the factor portfolios is minimum-variance efficient (Chamberlain (1983); Grinblatt and Titman, 1987). The consumption CAPM implies that a maximum correlation portfolio for consumption growth is minimum-variance
mimicking portfolios. Classical tests reject the model if the factor portfolio lies significantly inside a sample minimum variance boundary of factors and test assets (See Gibbons (1982); Stambaugh (1982); MacKinlay (1987); or GRS (1989).)

Let the N test assets’ excess returns be \( r \). Classical tests compare a maximum squared Sharpe ratio formed from the test assets and the factors, \( S^2(r,f) \), to one using only the factors, \( S^2(f) \). If the two squared Sharpe ratios are equal, the factor portfolio is efficient and the vector of the f-model’s average pricing errors or alphas (\( \alpha \)) for the test assets is zero. The alphas are the regression intercepts when the test assets’ excess returns, \( r \), are regressed over time on those of the benchmark factors, \( f \).

The difference between \( S^2(r,f) \) and \( S^2(f) \) is the squared Sharpe ratio of the Optimal Orthogonal Portfolio (OOP), a quadratic form in the f-model’s alphas for the test assets.

\[
(1) \quad \alpha'\Sigma(u)^{-1}\alpha = S^2(r,f) - S^2(f),
\]

where \( \Sigma(u) \) is the covariance matrix of the residuals of the test asset excess returns, \( r \), regressed over time, on the f-model’s factors. The OOP has the maximum mispricing using the f-model’s factors,\(^2\) Equation (1) is called a “law of conservation of squared Sharpe ratios” by Ferson efficient (Breeden, 1979). A stochastic discount factor model implies that a maximum correlation portfolio for the stochastic discount factor is minimum-variance efficient (Hansen and Richard, 1987).

\(^2\) The OOP maximizes its squared alpha divided by its residual variance, thus \( \arg \min_x x'\Sigma(u)x \) for a given \( x'\alpha \), with solution \( x=(\lambda/2)\Sigma(u)^{-1}\alpha \), where \( \lambda \) is the Lagrange multiplier. The maximized squared Sharpe ratio is \( (x'\alpha)^2/ x'\Sigma(u)x = \alpha'\Sigma(u)^{-1}\alpha \).
(2019). The maximum squared Sharpe ratio in all of the assets, $S^2(r,f)$ is equal to that in the tested factor portfolios, $S^2(f)$, plus that of the OOP.

A quadratic form in the alphas using the covariance matrix of the alphas in place of $\Sigma(u)$ is equal to $[S^2(r,f) - S^2(f)]/[1 + S^2(f)]$. Multiplying by a degrees-of-freedom adjustment $[(T-N-K)/N]$ produces the exact F test of GRS (1989). This paper shows how to conduct tests for factor-model efficiency that accommodate dynamic trading.

Barillas and Shanken (2017) consider the comparison of models with different factors, say $f_1$ and $f_2$, which might overlap. A model is better when the quadratic form in the test assets’ alphas using its factors is smaller. For example, model 1 is better than model 2 if $S^2(r,f_1,f_2) - S^2(f_1) < S^2(r,f_1,f_2) - S^2(f_2)$. Equivalently, model 1 is better if the maximum squared Sharpe ratio of its factors is larger: $S^2(f_1) > S^2(f_2)$. This paper shows how to conduct such direct factor model comparisons in terms of squared Sharpe ratios when there is dynamic trading.

A. Tests with Dynamic Trading

Consider a stochastic discount factor model:

\begin{equation}
E(m_{t+1} R_{t+1} \mid Z_t) = \mathbf{1},
\end{equation}

where $R$ represents the gross (one plus the rate of) return on the test assets, $m$ is the stochastic discount factor (SDF), $Z_t$ is the conditioning information, and $\mathbf{1}$ is an n-vector of ones. Linear factor models assume that the SDF is linear in the factors. When the SDF is linear in the factors, a linear combination of the factor portfolios is minimum variance efficient (Ferson (1995), Cochrane (1996)). This motivates tests comparing the squared Sharpe ratios of factor portfolios with those of test assets. This paper shows how to conduct such tests when there is dynamic trading. The conditioning information $Z_t$ is measured using lagged instruments, and portfolio
weights may vary as a function of these instruments. We use monthly data so the weights can vary monthly.

In principle, one could use conditional moments for the tests, as in early studies of conditional asset pricing (e.g., Hansen and Hodrick (1983), Gibbons and Ferson (1985), Harvey (1989)). These tests impose strong restrictions on the forms of the conditional first and second moments of the factors and test asset returns. Instead, we follow Ferson and Siegel (2009), who derive implications of Equation (2) for the unconditional moments of dynamic trading portfolios. Thus, the tests are based on unconditional squared Sharpe ratios, and the model comparison logic of Barillas and Shanken (2017) applies using unconditional squared Sharpe ratios. As shown by Bekaert and Liu (2004) and Ferson and Siegel (2003), such tests are inherently robust to misspecification of the conditional moments used in their construction.

Ferson and Siegel (2009) use an implication of Equation (2) for unconditional moments:

\[ E[m_{t+1} w'(Z_t) R_{t+1}] = 1 \quad \forall w(Z_t): w'(Z_t) 1 = 1. \]

As an implication of Equation (2), Equation (3) is less general, but not by much.\(^4\)

The tests examine unconditional efficiency (UE), which we now define. A set of dynamically trading portfolio returns determines a mean-standard deviation frontier, as shown by Hansen and Richard (1987), depicting the unconditional means versus the unconditional standard

\[^3\] The chosen Z is not likely to be the full information set used by market participants, but equation (2) may be arrived at by applying iterated expectations to a version of the equation conditioned on a finer public information set. We must restrict ourselves to models that are testable, in the sense that \( m_{t+1} \) depends only on the observable data and parameters that we can estimate.

\[^4\] Equation (2) is equivalent to the unconditional expectation \( E\{m_{t+1} R_{t+1} f(Z_t)\} = E\{1 f(Z_t)\} \) holding for all bounded integral functions, \( f(.) \). Equation (3) restricts to portfolio weight functions that sum to 1.0. Thus, for example with Equation (3) it is not possible to expand the scale of the risky investments without borrowing or lending at the risk-free rate.
deviations. A portfolio \( R_p \) is defined to be UE with respect to the information \( Z_t \) when it is on this frontier. Thus, Equation (4) is satisfied (equivalently, there exists constants \( \gamma_0 \) and \( \gamma_1 \) such that Equation (5) is satisfied) for all \( x(Z_t) \) such that the weights sum to one almost surely and the relevant unconditional moments exist and are finite:

\[
\text{Equation (4)} \quad \text{Var}(R_{p,t+1}) \leq \text{Var}\left[x'(Z_t)R_{t+1}\right] \quad \text{if} \quad E\left(R_{p,t+1}\right) = E\left[x'(Z_t)R_{t+1}\right]
\]

\[
\text{Equation (5)} \quad E\left[x'(Z_t)R_{t+1}\right] = \gamma_0 + \gamma_1 \text{Cov}\left[x'(Z_t)R_{t+1}, R_{p,t+1}\right].
\]

Equation (4) states that \( R_{p,t+1} \) is on the minimum variance boundary. Equation (5), shown by Hansen and Richard (1987), states that the expected return - covariance relation from Fama (1973) and Roll (1977) holds for the unconditional moments of dynamic portfolios. In Equation (5), \( \gamma_0 \) and \( \gamma_1 \) are fixed scalars that do not depend on the functions \( x(\cdot) \) or the realizations of \( Z_t \).

UE portfolios have proven useful in asset-pricing tests, in forming hedging portfolios and other portfolio management problems (e.g., Ahkbar, Devraj, and Stremme (2012), Chiang (2015), and Ehsani and Linnainmaa (2020)). Siegel (2021) provides a review. We develop asymptotic standard errors and bias correction for the squared Sharpe ratios of UE portfolios.

Ferson and Siegel (2009) show that, given an SDF such that equation (3) is satisfied, a portfolio with maximum correlation to \( m \) with respect to \( Z \) must be UE. The maximum correlation is among all portfolios of the test assets that may trade dynamically using the information, \( Z^5 \) This result provides the foundation for Sharpe ratio difference tests with dynamic

---

5 Formally, a portfolio \( R_p \) is maximum correlation for a random variable, \( m \), with respect to \( Z \), if:

\[
\rho^2(R_p,m) \geq \rho^2[w'(Z)R,m] \quad \forall w(Z): w'(Z)1 = 1,
\]
trading. To implement the tests with dynamic trading, we use closed-form solutions for UE portfolio weights, which exist for three relevant cases.\(^6\)

Ferson and Siegel (2015) generalize the law of conservation of squared Sharpe ratios (1) for dynamic trading. The dynamic OOP trades as a function of the lagged information to maximize the square of its unconditional alpha on a benchmark factor model divided by its unconditional variance of return\(^7\). A version of Equation (1) holds, with the unconditional squared Sharpe ratio of the dynamic OOP on the left hand side. The maximum unconditional squared Sharpe ratio of portfolios of the test assets and the factors that trade with the information, \(S_{ue}^2(r,f)\), replaces \(S^2(r,f)\).

The maximum squared Sharpe ratio of the factor benchmark with dynamic trading depends on the model. With dynamic trading using lagged instruments \(Z\), there are three ways to specify a linear factor model for the SDF (suppressing the time subscripts):

(i) \(m = a + b'f\), with fixed \((a,b)\),

(ii) \(m = a + b(Z)'f\), with fixed \(a\), and

(iii) \(m = a(Z) + b(Z)'f\).

\[\rho^2(.,.)\] is the squared unconditional correlation coefficient and we restrict to functions \(w(.)\) for which the correlation exists. Ferson, Siegel, and Xu (2006) present closed-form solutions for maximum correlation portfolios with respect to \(Z\).

\(^6\) In the first case there is a fixed risk-free rate, and in the second case there is no risk-free asset. Solutions are provided by Ferson and Siegel (2001). In the third case, there is a conditionally time-varying risk-free asset whose return \(R_{f+1} = R_f(Z_t)\) is measureable and known as part of the information set \(Z\), so that \(Var(R_{f+1}|Z_t) = 0\), but which is unconditionally risky in the sense that \(Var(R_{f+1}) > 0\). The solution is provided by Ferson and Siegel (2015) and Penaranda (2016). The same solution can be obtained using the expression from Ferson and Siegel (2001) for no risk-free asset, applied to the expanded return vector \((R_{t+1}, R_{f+1})\) (see the Internet Appendix, Section 3). Penaranda (2016) also describes residual efficiency, where the expected conditional variance is minimized for a given unconditional mean return.

\(^7\) There may be time-varying conditional alphas in this setting, but an unconditional regression is used to define the unconditional alphas.
In cases (ii) and (iii) the SDF is conditionally linear in the factors. There can be an error term uncorrelated with the test assets in each of these equations for $m$.

**Case (i):** If $m = a + b^\prime f$ with fixed (a,b) parameters then $b^\prime f$ is maximum correlation to $m$ and therefore is UE in the test asset returns $r$. Asking the model to price the excess returns of its own factors, $E(mf)=0$ identifies $b= -a E(f^\prime)^{-1}E(f)$ as a fixed minimum variance efficient portfolio weight for the factors. This motivates testing the hypothesis that $S^2_{fix}(f) = S^2_{ue}(r)$.

**Case (ii):** The portfolio $b(Z)^\prime f$ is maximum correlation to $m$ when $a$ is a constant, and therefore it must be UE. $E(mf|Z)=0$ identifies $b(Z) = -a E(f^\prime|Z)^{-1} E(f|Z)$ as UE weights for the factors. This motivates testing the hypothesis that $S^2_{ue}(f) = S^2_{ue}(r)$.

**Case (iii):** We call this a “fully conditional model.” The maximum correlation portfolio to $a(Z) + b(Z)^\prime f$, with respect to $Z$, should be UE in $r$. $E(mf|Z)=0$ identifies $b(Z) = -a(Z) E(f^\prime|Z)^{-1} E(f|Z)$ and $b(Z)^\prime f$ is conditionally minimum variance in $f$ but need not be UE. We can find $a(Z)$ assuming a conditionally risk-free rate that is known given $Z_t$ from $E\{mR_t|Z_t\}=1$, as $a(Z) = (1/R^\prime_t)[1-E(f^\prime|Z)^{-1} E(f|Z)]$. Using this to construct the maximum correlation portfolio to $m$, we can test the hypothesis that is is UE by comparing squared Sharpe ratios.

### III. Main Results

#### A. Asymptotic Variances

The asymptotic variance of an estimator $\hat{\theta}$, that depends on the conditional mean vector and covariance matrix estimators, such as a squared Sharpe ratio, may be characterized in terms of two canonical matrices $C (L \times N)$ and $D (N \times N)$ that capture the sensitivity of the estimator to uncertainty in the conditional means and in the conditional covariance respectively. We first present a general theorem, and then derive special cases. The proofs are provided in the Internet Appendix.
Our analysis is based on the following assumptions. The conditional expected returns are described by a linear regression:

\[ R_t = \delta' Z_{t-1} + \epsilon_t. \]

The return vector for \( N \) assets at time-\( t \) is \( R_t \) and the \( L \) lagged instruments (including a constant) are \( Z_{t-1} \) for \( t = 1, \ldots, T \). The coefficient \( \delta \) is an \( L \times N \) matrix of regression coefficients. The conditional mean returns at time-\( t \) are \( \delta' Z_{t-1} \). The regression errors \( \epsilon_t \) are independent and identically distributed (iid) \( N \times 1 \) vectors with mean zero and nonsingular conditional covariance matrix, \( V = E \left( (R_t - \delta' Z_{t-1})(R_t - \delta' Z_{t-1})' | Z_{t-1} \right) = E (\epsilon_t \epsilon_t' | Z_{t-1}) \).

We estimate the OLS regression (6) and express:

\[ \hat{\delta} - \delta = A^{-1} \left( \frac{1}{T} \sum_{t=1}^{T} Z_{t-1} \epsilon_t' \right) = O_p \left( 1 / \sqrt{T} \right). \]

where the \( L \times L \) matrix \( A = E(Z'Z/T) \) is assumed to be nonsingular. We define the estimated covariance matrix as \( \hat{V} \equiv \frac{1}{T} \sum_{t=1}^{T} (R_t - \hat{\delta}' Z_{t-1})(R_t - \hat{\delta}' Z_{t-1})' = \frac{1}{T} \sum_{t=1}^{T} \hat{\epsilon}_t \hat{\epsilon}_t' \), using the estimated residuals, \( \hat{\epsilon}_t \equiv R_t - \hat{\mu}_t = R_t - \hat{\delta}' Z_{t-1} \).

**Theorem I:** Consider a scalar estimator of the form

\[ \hat{\theta} = \theta + \sum_{t=1}^{T} C_t' (\hat{\mu}_t - \mu_t) + tr \left[ D (\hat{V} - V) \right] + O_p \left( 1 / T \right) \]

where \( C_1, \ldots, C_T \) are \( N \times 1 \) vectors that depend on \{\( Z_t \)\} and \( D \) is a fixed \( N \times N \) matrix. Defining the \( L \times N \) matrix \( C \equiv A^{-1} \sum_{t=1}^{T} Z_{t-1} C_t' \), where \( A = E(Z'Z/T) \). The unconditional asymptotic variance of \( \hat{\theta} \) is estimated as:
where $\hat{C}$ and $\hat{D}$ are consistent estimates for the canonical matrices $C$ and $D$ respectively, and

$$\bar{Z} = \frac{1}{T} \sum_{t=1}^{T} Z_{t-1}.$$ Let $\psi_t = Z_{t-1}' C \varepsilon_t + \varepsilon_t' D \varepsilon_t$. The unconditional variance of $\hat{\theta}$ may be estimated as the sample variance of the $\psi_t$:

\begin{equation}
\sqrt{T} \{ (1/T) \Sigma_t \psi_t^2 - [(1/T) \Sigma_t \psi_t]^2 \},
\end{equation}

where consistent estimates of $C$ and $D$ can replace the true parameter values in $\psi_t$.

Squared Sharpe ratios are a natural special case of Theorem I as their sampling errors depend on the sampling errors in a vector of mean returns and a covariance matrix. The asymptotic variance of a function of two estimators may be found by using estimates $\hat{C}$ and $\hat{D}$ of the canonical matrices associated with each estimator. For example, the asymptotic variance of the difference between two squared Sharpe ratios, $\hat{\theta} = \hat{S}^2 (r)$ and $\hat{\theta}' = \hat{S}^2 (r')$ has canonical matrices equal to the differences in the two canonical matrices, $\hat{C} - C'$ and $\hat{D} - D'$. Using the differences allows for dependence. For example, the $(C - C')' V (C - C') A$ in the first term of (8) accounts for the covariance due to common dependence on the estimation error in the conditional means. The third term of (8) involves higher comoments. The asymptotic variance of the ratio $\hat{\theta} / \hat{\theta}'$ has canonical matrices $\frac{\hat{\theta}' \hat{C} - \hat{\theta} \hat{C}'}{\hat{\theta}'^2}$ and $\frac{\hat{\theta}' \hat{D} - \hat{\theta} \hat{D}'}{\hat{\theta}'^2}$. The asymptotic variance of $\frac{\hat{\theta} - \hat{\theta}'}{1 + \hat{\theta}'}$ has canonical matrices $\left[ (1 + \hat{\theta}') \hat{C} - (1 + \hat{\theta}) \hat{C}' \right] / (1 + \hat{\theta}')^2$ and $\left[ (1 + \hat{\theta}') \hat{D} - (1 + \hat{\theta}) \hat{D}' \right] / (1 + \hat{\theta}')^2$. 

\begin{equation}
AVAR(\hat{\theta}) \equiv tr \left( \hat{C} \hat{V} \hat{C}' A \right) - \left[ tr \left( \hat{D} \hat{V} \right) \right]^2 + \frac{1}{T} \sum_{t=1}^{T} \left( 2 \bar{Z}' \hat{\varepsilon}_t + \hat{\varepsilon}_t' \hat{D} \hat{\varepsilon}_t \right) \left( \hat{\varepsilon}_t' \hat{D} \hat{\varepsilon}_t \right)
\end{equation}
Our main results assume a constant covariance matrix, $V$, but stock returns are known to be conditionally heteroskedastic through time. The Internet Appendix presents an extension of the Theorem I that allows for heteroscedasticity. But the results will depend on the particular model used for the time-varying covariances, and we leave the full development of this case for future research.

Tests such as case (i) above require comparing Sharpe ratios for fixed-weight portfolios with those that use the information. Theorem I can be applied to such cases, as described in the Appendix as Corollary I.

**B. Distributions of the Test Statistics**

Our suggestion is to test models with a “t-ratio,” dividing the bias-adjusted squared Sharpe ratio or difference by its asymptotic standard error. We first argue that the estimated squared Sharpe ratios are consistent and asymptotically normal. The argument uses the Generalized Method of Moments (Hansen, 1982), similar to BKRS (2019) for the fixed-weight case.

Define the moment conditions $g_{1t} = \text{Vec}(r_t - \delta'Z_t Z_t^{-1})$ and $g_{2t} = \text{Vech}(\varepsilon_t'\varepsilon_t - V)$. Stack them into $g_t = (g_{1t}', g_{2t}')$ and let $g = (1/T)\Sigma_t g_t$. For any quadratic form $E(g)'W E(g)$, the GMM estimates for the parameters $\phi = (\text{Vec}(\delta)'\text{Vech}(V)'$ are found by setting $E(g)'W E(g) = 0$, as the problem is exactly identified. The GMM estimates are the OLS estimates, and they are the same when a consistent estimate for the asymptotic covariance matrix of $g$ is used for the weighting matrix, $W$. Hansen’s (1982) Theorems 2.1 and 3.1 show that the estimates of $\phi$ are consistent and asymptotically normal on the assumptions that $\{g_t\}$ is strictly stationary and ergodic, $g$ has continuous derivatives with respect to the parameters in a neighborhood of the true values, and the parameters lie in a compact set. Given these results we argue that the estimated squared Sharpe ratios less their true values are asymptotically normal. We can express the squared
Sharpe ratio or difference as a function $\theta = \theta(\delta, V)$ and take the first-order derivatives to find $Avar(\theta) \approx (\partial \theta / \partial \phi) Avar(\phi) (\partial \theta / \partial \phi)'$. Our Theorem I presents asymptotically equivalent expressions for standard errors.

A quadratic form in the estimated squared Sharpe ratio differences, using a consistent estimator of their variance as the inverse matrix in the form, is asymptotically distributed as a Chi squared random variable. Since the squared Sharpe ratio or difference is a scalar, the square of the t-ratio that we propose is asymptotically distributed as $\chi^2(1)$. The absolute value of the t-ratio is therefore asymptotically distributed as a $\chi(1)$, or half-normal random variable. In most applications positive Sharpe ratio differences are examined, so our simulations evaluate the absolute t-ratios relative to a Chi distribution with one degree of freedom. The critical values of this density are similar to the standard rules of thumb used with t-ratios. For example, the critical value for 5% significance is 1.96 and for 10% significance it is 2.33.

We use T-consistent estimates for a finite number of assets, $N$. It is therefore important to understand how the estimators perform when confronted in practice with relatively large values of $N$. Factor model comparisons are typically conducted using portfolios, so we present simulations using up to 99 portfolios.

C. Asymptotic Variances for UE Portfolio Squared Sharpe Ratios

We start with results from Ferson and Siegel (2001, Theorem 3), who derive the portfolio weights of the UE portfolio when there is no risk-free asset. The portfolio has target unconditional mean $\mu_p$ and the weight is:

$$w_i' = \frac{1' \Lambda_t}{1' \Lambda_t 1} + \frac{\mu_p - \alpha_3}{\alpha_3} \mu_i' \left( \Lambda_t - \frac{\Lambda_t 11' \Lambda_t}{1' \Lambda_t 1} \right)$$
where: \[ \Lambda_t \equiv (\mu, \mu_t' + V)^{-1}, \]
and the efficient set constants \( \alpha_1, \alpha_2, \) and \( \alpha_3 \) are defined in Ferson and Siegel (2001) and in our Internet Appendix. The portfolio unconditional variance is \[ \sigma^2_p = \left( \alpha_1 + \frac{\alpha_2}{\alpha_3} \right) - \frac{2\alpha_2}{\alpha_3} \mu_p + \frac{1 - \alpha_3}{\alpha_3} \mu_p^2. \]

The estimators based on \( \hat{\mu}_t \) and \( \hat{V} \) are denoted \( \hat{\Lambda}_t, \hat{w}_t, \hat{\alpha}_1, \hat{\alpha}_2, \) and \( \hat{\alpha}_3. \) Our goal is to estimate the asymptotic variance of the estimated maximized squared Sharpe Ratio of the portfolio.

**Corollary I**: The asymptotic variance of the squared Sharpe Ratio of the UE portfolio with the weights in Equation (10), when \( \mu_p \) corresponds to the zero-beta rate \( \varphi, \) may be obtained using Theorem I with canonical matrices: \(^8\)

\[
C = -\left[ \alpha_2 - \varphi (1 - \alpha_3) \right] C_{\alpha_1} + 2 \left( \alpha_1 - \varphi \alpha_2 \right) \left[ \alpha_2 - \varphi (1 - \alpha_3) \right] C_{\alpha_2} + \left( \alpha_1 - \varphi \alpha_2 \right)^2 C_{\alpha_3} \]

(11)

\[
D = -\left[ \alpha_2 - \varphi (1 - \alpha_3) \right] D_{\alpha_1} + 2 \left( \alpha_1 - \varphi \alpha_2 \right) \left[ \alpha_2 - \varphi (1 - \alpha_3) \right] D_{\alpha_2} + \left( \alpha_1 - \varphi \alpha_2 \right)^2 D_{\alpha_3} \]

where the component matrices \( C_{\alpha_1}, D_{\alpha_1}, C_{\alpha_2}, D_{\alpha_2}, C_{\alpha_3}, \) and \( D_{\alpha_3} \) and their consistent estimates are provided with the proof in the Internet Appendix.

Corollary I may be used to compare models with dynamic trading by using the differences in the portfolios’ canonical matrices. For example, in the nested tests where model 1 adds factors

\(^8\) When we say that \( \mu_p \) corresponds to the zero-beta rate \( \varphi, \) we mean that a tangent line drawn to the minimum standard deviation boundary at the portfolio with mean \( \mu_p \) intersects the y-axis at the point \( \varphi. \) For a given boundary only one of the two parameters \( (\mu_p, \varphi) \) need be specified. The variance of a minimum variance efficient portfolio is a quadratic function of its mean: \( \sigma_p^2 = a - 2b\mu_p + c\mu_p^2. \) The slope of the line from a zero beta rate, \( \varphi, \) to this boundary, \( \partial \mu_p/\partial \sigma_p = (\mu_p - \varphi)/\sigma_p. \) Thus, the target mean is related to the zero beta rate as: \( \mu_p = (a - b\varphi)/(b - c\varphi). \)
to those of model 2, the matrices $C_i$ and $D_i$ are formed using the larger model 1 factors, and the model 2 matrices are computed by filling the positions for the additional factors with zeros.

D. Zero-Beta Rates

The analyses of BS (2017), FF (2018), and BKRS (2020), and others assume that the average short-term Treasury rate is the zero beta rate. However, models and evidence often imply zero beta rates in excess of a short-term Treasury return (e.g. Black, Jensen, and Scholes (1972), Frazzini and Pedersen (2014), Lewellen, Nagel, and Shanken (2010)). The difference between the zero beta rate and the Treasury rate can reflect the costs of borrowing against the underlying assets (Black (1972), Frazzini and Pedersen (2014)). Because the zero-beta rate is the expected return on assets that have zero covariance with the factors in a model, if there are missing priced factors the zero beta portfolio may embed a premium because of its correlation with the missing priced factors.

We consider three alternative treatments for the zero beta rate. The first is to work with returns in excess of a Treasury bill, as in most of the recent work on Sharpe ratio comparisons. The mean value of the Treasury bill return is the expected zero beta rate. The second approach is to allow independent dynamic trading in the Treasury bill; viewed as a time-varying risk-free asset return, which is conditionally known at the beginning of the period. This approach is implemented using the solutions provided by Ferson and Siegel (2015). Our third approach is to estimate the zero beta rate while assuming that no risk-free asset exists.

We estimate a zero beta parameter following Kandel (1986), who worked in a normal, maximum likelihood setting. Kandel derived the zero beta rate that maximizes the log likelihood ratio, the difference between the log likelihood function that imposes the null hypothesis and the
log likelihood for the unrestricted data. The null hypothesis is $S^2(r,f)=S^2(f)$, so the optimal zero beta rate minimizes the difference between the two maximum squared Sharpe ratios. The following proposition provides the solution, exploiting the fact that the maximum squared Sharpe ratio of any portfolio is a quadratic function of the zero beta rate: $\theta(\phi) = a - 2b\phi + c\phi^2$

where the coefficients ($a$, $b$, and $c$) depend on the situation.

**Proposition I:** The zero beta rate $\phi$ that minimizes the difference between the maximum squared Sharpe ratio of a portfolio with squared Sharpe ratio $\theta(\phi) = a - 2b\phi + c\phi^2$ and that of another with maximum squared Sharpe ratio $\theta^*(\phi) = a^* - 2b^*\phi + c^*\phi^2$, is found by minimizing $(a^* - a) - 2(b - b^*)\phi + (c - c^*)(\phi^2)$ and the solution is:

(12) \[ \phi = \frac{b - b^*}{c - c^*}. \]

The zero beta rate in the case of a normalized difference, like the test of GRS (1989), should minimize: \[ \frac{1 + \theta(\phi)}{1 + \theta^*(\phi)} = \frac{1 + a - 2b\phi + c\phi^2}{1 + a^* - 2b^*\phi + c^*\phi^2}. \]

The first-order condition delivers the quadratic expression:

\[ (b^* - b) + (ab^* - a^*b) + (-2(a + 1)c^* + 2(a^* + 1)c)\phi + (-3bc^* + 3b^*c)\phi^2 = 0. \]

There are two solutions to this equation, one representing the local maximum and one the minimum. The consistent estimates are found by substituting consistent estimates of the $(a,b,c)$ parameters into Equation (12) or the solution to the quadratic equation.

When the optimal zero beta rate from Proposition I is used in the comparison of squared Sharpe ratios, we accommodate the estimation error in the common zero beta rate as another special case of Theorem I, given in Corollary II.
Corollary II: Two maximum squared Sharpe ratios can be written as $\theta = a - 2b\phi + c\phi^2$ and $\theta^* = a^* - 2b^*\phi + c^*\phi^2$. By proposition I the optimum zero-beta rate is $\phi = \frac{b-b^*}{c-c^*}$. The difference between two squared Sharpe ratios is $\theta - \theta^*$. Assume the estimator for $a$, $\hat{a}$, can be expanded as in the Theorem I, with canonical matrices $C_a$ and $D_a$. Make the same assumption for $\hat{b}$, $\hat{c}$, $\hat{a}^*$, $\hat{b}^*$, and $\hat{c}^*$. Then, the estimated value of the difference between two squared Sharpe ratios with a common estimated optimal zero-beta rate can be written as in the Theorem I, with canonical matrices:

\begin{align}
\text{(13)} & \quad C = C_a - C_a^* - 2\phi(C_b - C_b^*) - 2(b - b^*)C_c + 2\phi(C_c - C_c^*) + 2\phi(c - c^*)C_\phi, \\
\text{(14)} & \quad D = D_a - D_a^* - 2\phi(D_b - D_b^*) - 2(b - b^*)D_c + 2\phi(D_c - D_c^*) + 2\phi(c - c^*)D_\phi,
\end{align}

where

\begin{align*}
C_\phi &= \frac{1}{c-c^*} C_b - \frac{1}{c-c^*} C_b^* - \frac{b-b^*}{(c-c^*)^2} C_c + \frac{b-b^*}{(c-c^*)^2} C_c^*, \\
D_\phi &= \frac{1}{c-c^*} D_b - \frac{1}{c-c^*} D_b^* - \frac{b-b^*}{(c-c^*)^2} D_c + \frac{b-b^*}{(c-c^*)^2} D_c^*.
\end{align*}

The expressions are evaluated at consistent estimates of $(a,b,c)$ and $(a^*,b^*,c^*)$. The Internet Appendix presents the matrices for cases where one of the portfolios is a maximum correlation portfolio with respect to $Z$, as in models with nontraded factors.

E. Bias Correction

Jobson and Korkie (1980), provide a bias adjustment for sample squared Sharpe ratios based on the assumption that returns are normally distributed and the optimal portfolios have fixed weights. This adjustment is used by Ferson and Siegel (2003) and BKRS(2020). The bias-corrected estimate, improving the biased estimate $\hat{\theta}$ is given by:
\[ \hat{\theta}^* = \hat{\theta} \left( \frac{T - N - 2}{T} \right) - \frac{N}{T} \]

Where \( N \) is the number of assets and \( T \) is the number of time-series observations.

Ferson and Siegel (2003) find that the bias adjustment in (15) does not control the bias very well for dynamic portfolios, so we evaluate alternative bias adjustments. One is based on the noncentral Chi-square distribution, assuming that the sample covariance matrix is at its probability limit. A second is based on the noncentral F distribution, which considers estimation error in the covariance matrix. A third is based on the odd-month, even-month approach of Jegadeesh, Noh, Pukthuanthong, Roll, and Wang (2019), which attempts to control for correlated errors. A fourth, and it turns out the best approach, is based on a second-order Taylor expansion.

Our preferred bias adjustment is based on the exact expectation of the second-order Taylor series expansion of the estimate minus its true value. The results are T-consistent when consistently estimated values are substituted for the unknown parameters. The bias correction is expressed in a proposition, where we define the squared Sharpe ratio in terms of the coefficients \( (\alpha_1, \alpha_2, \alpha_3) \) as in Proposition I.

**Proposition II**: The approximate bias of the estimated maximized squared Sharpe Ratio at zero-beta rate \( \psi \): 

\[ \hat{S}^2_{\psi} = \frac{\hat{\alpha}_1 - 2\psi \hat{\alpha}_2 + \psi^2 \hat{\alpha}_3}{\hat{\alpha}_1 \hat{\alpha}_3 - \hat{\alpha}_2^2} - 1 \]

with respect to the true (but unknown) maximized squared Sharpe Ratio 

\[ S^2_{\psi} = \frac{\alpha_1 - 2\psi \alpha_2 + \psi^2 \alpha_3}{\alpha_1 \alpha_3 - \alpha_2^2} - 1 \]

may be expressed using the expectation of its second-order Taylor Series expansion:

\[ \text{Bias} = E \left( \hat{S}^2_{\psi} - S^2_{\psi} \right) \approx \left( \sum_{i=1}^{3} E \left( \hat{\alpha}_i - \alpha_i \right) \frac{\partial S^2_{\psi}}{\partial \alpha_i} \right) + \left( \sum_{i,j=1}^{3} \frac{E \left[ (\hat{\alpha}_i - \alpha_i)(\hat{\alpha}_j - \alpha_j) \right]}{2} \frac{\partial^2 S^2_{\psi}}{\partial \alpha_i \partial \alpha_j} \right) \]
As in Siegel and Woodgate (2007a,b), we use the method of statistical differentials to find Taylor-series approximations to expectations of random variables. The partial derivatives in Proposition II are presented in the Internet Appendix. Note that the \( \hat{\alpha}_i \) are not necessarily unbiased, so their expectations are necessary in the expression (16).

**IV. The Data**

We follow Cooper and Maio (2019) in the selection of six lagged information variables. These are a short-term Treasury bill rate, a value spread (Cohen, Polk, and Vuolteenaho (2003)), a measure of stock return dispersion following Stivers and Sun (2010), net equity expansion following Boudoukh, Michaely and Richardson (2007), and the investment to capital ratio following Cochrane (1991). We also present some results using a set of “classical” lagged instruments following Fama and French (1989) and a more “modern” set of instruments following Goyal, Welch, and Zafirov (2021).

The lagged conditioning variables in much of the literature are highly persistent. We assume that they are stationary, but if they are unit root processes, the distribution of the sample squared Sharpe ratios will be nonstandard (e.g., Phillips, 2014). We address this by following the suggestion of Ferson, Sarkissian, and Simin (2003). We subtract a 12-month trailing average from each of the lagged instruments to stochastically detrend them when the first-order autocorrelation exceeds 0.95. All of the autocorrelations of the stochastically detrended series

\[ \text{\textsuperscript{9}} \]

One exception is the investment-to-capital ratio, which is available quarterly. Cooper and Maio (2019) fill in for monthly data, assuming a constant value for the months within a quarter. This produces high autocorrelation which the 12-month moving average does not correct. But a 6-month moving average works for this series.
are well below values that Ferson, Sarkissian, and Simin find raise concern over spurious regression bias.

Standard test portfolio returns are monthly data from Kenneth French’s data library at Dartmouth. Individual stocks sorted in two or more dimensions to form cross-sections of portfolio returns. We use the 25 size × value portfolios, the 25 investment × profitability portfolios, 32 size × investment × profitability triple-sorted portfolios, and 49 industry portfolios.

We compare a number of popular factor models. The factors include the CRSP value-weighted stock market index, the Fama-French (1996) three factors (FF3), and the Fama and French (2015) five factors (FF5). We also examine the four-factor (Q4) model of Hou et al. (2015), using data on the investment factors from Lu Zhang, and the Q5 model of Hou et al. (2021), using data on investment growth. In the Internet Appendix, we examine non-traded factors, following Chen et al. (1986).

[Insert Table 1 about here]

The summary statistics for the factors and conditioning information are presented in Table 1. Returns are measured in excess of a three-month Treasury bill. The first-order autocorrelations of the factors are 0.33 or below. Even with these small values, Ehsani and Linnainmaa (2020) find that dynamic trading using lagged factors as the information can materially increase Sharpe ratios. This is consistent with the logic of Campbell (1996) and Cooper and Maio (2019).

V. Simulations

We use the parametric bootstrap to evaluate the corrections for finite sample bias in estimated Sharpe ratios, the accuracy of our asymptotic standard errors, and the finite sample distributions of the “t-ratios.” We model the Zs in the simulations as a first-order autoregressive process to capture their persistence.
A. Simulation Methods

Assume that the portfolio returns and the factors in the model follow equations (17)-(18):

\[(17)\quad R_t = \delta' Z_{t-1} + \epsilon_t,\]

\[(18)\quad f_t = b_F' Z_{t-1} + \epsilon_{ft},\]

and the conditioning information (without the constant) \(Z\) follows an AR(1) process:

\[(19)\quad Z_t = \delta_{zt} + \delta_{zt-1} + \epsilon_{z_t}.\]

We estimate the coefficients in the original data and calculate the regression residuals. We keep the coefficients as “true” parameters in our simulations. At each date in a simulation trial, we randomly choose a calendar date from the real data, and select the regression residuals from equations (17)-(19) as a vector for that date to preserve the correlations across the shocks. We take the sample average value of each conditioning variable in the data as its starting value, and build up the simulated instruments series recursively using the residuals and the regression coefficients of Equation (19). We discard the first 500 simulated samples to wash out the initial conditions.

B. Simulation Results

1. Evaluating Bias Adjustments

The estimators are consistent, converging in probability to the true values as the number of time-series observations, \(T\), grows. We evaluate finite sample bias relative to a “true” squared Sharpe ratio where the number of time series observations is \(T*1000\), and \(T=743\) or 587.

Informal experiments suggest that 1000*\(T\) is larger than required to estimate the probability limits. In each of 5,000 artificial samples with the same length as the original data samples, we estimate the bias-adjusted squared Sharpe ratios. The expected adjusted squared Sharpe ratio in finite samples is computed as the average value across the 5,000 trials.
Panel A of Table 2 evaluates bias adjustment for fixed-weight portfolios of test assets and factors. The unadjusted squared Sharpe ratios are larger than the true values, reflecting the finite sample bias. The Jobson and Korkie (1980) adjustment works well when there is no dynamic trading. After adjustment, the remaining bias is 5% or less of the true value. The bottom rows show results for N=25, 49, and 99 portfolios. The N=25 portfolios are the 5×5 size × book/market sorts, the N=49 are industry portfolios and the N=99 combine the first two sets with 25 investment × profitability portfolios. The bias shows no obvious relation to N.

Panel B of Table 2 looks at UE portfolio squared Sharpe ratios. Comparing the left-hand columns of the panels A and B of Table 2 shows the impact of dynamic trading on the “true” maximum squared Sharpe ratios. The impact is large for $R_m$, the FF6 squared Sharpe ratio increases by more than 50% with dynamic trading, and the Q5 by about 11%. The impact of dynamic trading on the standard portfolio designs is larger still, and the maximum squared Sharpe ratios of the FF25 and the 49 industry portfolios more than double.

Panel B of Table 2 addresses the various adjustments for finite sample bias with dynamic trading. The unadjusted squared Sharpe ratios can have a large bias, increasing when there are more assets or factors. The Q5 model, which delivers the highest Sharpe ratios among the factor models, has the smallest bias. While all of the bias correction methods reduce the finite sample bias, the direct expansion approach is clearly the most accurate. The remaining bias in the adjusted ratios is less than 3.2% of the true value in each case, and the percentage bias shows no obvious relation to N. The direct expansion bias adjustment works as well on the UE portfolios as the JK (1980) adjustment does on the fixed-weight portfolios. While not shown in the table,
we also find that the direct expansion method works well on fixed-weight portfolios, but its calculation is not as simple as the JK adjustment in such cases.

2. **Asymptotic Standard Errors**

We examine the accuracy of our asymptotic standard errors for squared Sharpe ratios and their differences. The standard deviations of the ratios estimated in finite samples, and bias adjusted with the second-order expansion method, are taken across 1,000 simulation trials. These empirical standard errors measure the sampling variability in the estimates. Good asymptotic standard errors should predict the empirical standard errors.

[Insert Table 3 about here]

The average value of the asymptotic standard errors across the simulation trials is shown as Average Asymptotic in Table 3. Panel A describes results for the levels of squared Sharpe ratios. The asymptotic standard errors present an expected bias averaging about 7% of the empirical overall, but as high as 11% (fixed-weight FF3 model). The asymptotics tend to understate the sampling variability. In the case with fixed portfolio weights, our asymptotic standard errors (FSW) are very similar to the BKRS standard errors. The FSW standard errors are typically more accurate in the cases with dynamic trading than with fixed weights.

The rows in Table 3 for the fixed-weight efficient portfolios show that Theorem I may be used for the standard errors in this case. These simulations do reflect the randomness of the conditioning information.\(^\text{10}\) We see no strong evidence that the bias in the standard errors is larger when more portfolios are used.

---

\(^\text{10}\) To further explore the impact of the volatility of the conditioning information, we shut it off by using the same values of the conditioning variables at each simulation trial. Otherwise, the simulations are the
Panels B–D of Table 3 examine standard errors for Sharpe ratio differences, as used in tests of the models. The average asymptotic values are within 10% of the empirical standard errors in most cases and there is no indication that the performance degrades when more test asset portfolios are used. The accuracy of the standard errors is somewhat worse when combinations of factors are compared directly in Panel B, where the asymptotic standard error overstates the sampling variability.

3. T ratios

We form “t-ratios,” dividing the bias-adjusted squared Sharpe ratio or difference by its standard error. Table 4 evaluates the sampling distributions of the absolute t-ratios against their asymptotic distribution, a Chi distribution with one degree of freedom. The true values of the numerators are from simulations with 743,000 observations. Fractiles of the distribution of the t-ratios from the 1,000 simulation trials are shown. $\chi(1)$ are the critical values for the asymptotic distribution, which are very similar to the usual rules of thumb as can be seen in the table.

[Insert Table 4 about here]

Panel A of Table 4 presents critical t-ratios for the levels of maximum squared Sharpe ratios. Using the standard test portfolios (N=25, 49, or 99) the t-ratios appear reasonably well specified in the tails. When small numbers of factors are used, the performance is worse and the empirical critical values are too large. In the worst cases, however, a t-ratio of 2 is significant at the 10% level instead of the 5% level. This is reminiscent of results from GRS (1989) where the

______________________________

same as before. We find that the average FSW asymptotic and the empirical standard deviations are even closer to each other in this experiment.
Wald and Lagrange multiplier tests rejected the CAPM too often in various test portfolio designs. Here the CAPM displays the largest bias.

Panels B–D of Table 4 present results for squared Sharpe ratio differences. The bias remaining in the bias-adjusted squared Sharpe ratio levels may be offsetting in the differences. The table shows that the distributions of the sample t-ratios are close to the $\chi(1)$ in most cases. This supports the validity of a standard rule of thumb where a t-ratio of 2 is considered significant at the 5% level in a maximum squared Sharpe ratio difference test between a factor model and the standard test asset portfolios.

Panel E of Table 4 presents results for direct comparisons between the squared Sharpe ratios of two models’ factors. In this application the t-ratios are the least well specified and are expected to reject the hypothesis of no difference between the two models too often. The fixed-weight cases are worse than the UE cases, but the Q5 model produces better specified t-ratios.

VI. Factor Model Tests and Comparisons

Table 5 presents tests of factor model efficiency in various test portfolio designs. Returns are measured in excess of a three-month Treasury bill rate. Its average value, 0.39% per month, is the zero beta rate. In panel A, no conditioning information is used. We include a factor model from Chen et al. (CRR3, 1986) where mimicking portfolios replace the nontraded factors for unexpected inflation, expected inflation, and industrial production. Most of the bias-adjusted Sharpe ratio differences between test assets and factors are statistically significant (t-ratios of 2.2 or higher) except where the portfolios sort on profitability and the factor models feature investment, profitability or production factors (FF5, FF6, Q4, Q5, and CRR3). Using the value-weighted market index, the CAPM is rejected in all the portfolio designs, and all the models are rejected in the size x value portfolios.
Given that factor models with fixed weights are rejected in test portfolios with no dynamic trading, it is clear (and we do find) that factor models with fixed weights can be rejected in test portfolios with dynamic trading. Thus, case (i) above for linear factor models of the SDF with constant coefficients can be rejected. In Panel B of Table 5, dynamically trading factor models face dynamically trading test assets. The Sharpe ratio differences are larger than in Panel A, with large t-ratios in almost all cases. This is because dynamic trading increases the Sharpe ratios of the test assets more than factors. Thus, case (ii) for linear factor models of the SDF can be strongly rejected.

In Panel C of Table 5, the fully conditional factor models of case (iii) are tested, and rejected with t-ratios larger than 2.16 in every case. With the time-varying a(Z) term that does not depend on the factors, the models are less sensitive to the choice of factors, and the Sharpe ratios and test statistics are similar across the factor models.

The strong rejections of the models in Table 5 assume that the zero beta rate is the average Treasury bill rate of 0.39%. In Table 6 this assumption is removed, and the zero beta rate is estimated. Panel A uses fixed-weight portfolios. There are far fewer rejections of the models allowing the more general zero-beta rates. T-ratios larger than 2 only appear in the industry portfolio design. The CAPM produces both the largest t-ratio and the largest zero beta rate among the factor models, with zero beta rates of 0.8% to 1.5% per month. The zero beta rate can reflect borrowing costs or the average risk premiums for missing factors. It seems likely that the high zero beta rates for the CAPM reflect the impact of missing factors. For other models the zero beta rate is closer to the average bill rate.
In Panel B of Table 6, dynamic trading is used for both factors and test assets. The zero beta rate estimates are similar to those of Panel A, but we find more frequent rejections of the models as dynamic trading increases the Sharpe ratios for the test assets more than it does for the factors. All the models except CRR3 are rejected in the industry portfolio design.

A. Direct Factor Model Comparisons

Table 7 presents direct comparisons of the maximum squared Sharpe ratios for two models’ bias-adjusted factors. In the first column, the factors are held with fixed weights over time. The next three columns use dynamic trading with different choices of lagged instruments, described in the data section. The results are similar for the different instrument choices.

[Insert Table 7 about here]

Table 7 shows that dynamic trading helps the FF3 factors beat the market index, but makes less of a difference for the other model comparisons. The Sharpe ratio performance of the Q models is striking. The Q5 model outperforms every model to which it is compared. This includes a version of FF6, denoted FF6* in the table, where a monthly rebalanced HML factor is used.

The FF5 model beats FF3 in Table 7, but the FF6 model (with a momentum factor) only beats the FF5 when there is dynamic trading using the “modern” instruments. Fama and French (2018) equivocate on whether the FF6 beats the FF5 model. The Q4 model beats FF5 in a fixed-weight comparison, but the two are not significantly different with dynamic trading. Dynamic trading improves the FF5 factors more than the Q4 factors.

The tests in Table 7 are based on returns in excess of the Treasury bill. We conduct, but do not report in the tables, direct model comparisons where we estimate an optimal zero beta rate. In only one of the direct comparisons is the null hypothesis rejected that the two Sharpe ratios
are equal with dynamic trading. Unlike in tests using the standard test asset portfolios, the optimal zero beta rates vary widely over these model comparisons, but they are not very sensitive to the choice of instruments.

VII. Conclusions

This paper contributes to the literature on factor model tests and comparisons in three ways. First, we remove the limitation that optimal portfolios must have fixed weights over time, allowing dynamic trading as a function of lagged variables. Second, we develop finite sample bias adjustments for maximum squared Sharpe ratios. Third, we allow general and optimal zero beta rates in the models.

We develop asymptotic standard errors and propose simple “t-ratios” that our simulations show are reliable in most cases for realistic samples of stock portfolio returns and popular factor models. This allows inferences without simulation.

The impact of dynamic trading with the conditioning information is to raise the maximum squared Sharpe ratios, and to a larger extent for popular portfolios sorted on stock characteristics or industries than for portfolios of the models’ factors. Most of the popular factor models are thereby rejected, and the factors in the models appear far from efficient even when the factors are traded dynamically. This presents the awkward situation where the factor models’ Sharpe ratios appear “too large” from some perspectives, yet “too small” to be efficient in the standard portfolio designs.

In factor model efficiency tests using standard test portfolios, we find that the estimated zero beta rate is the largest in the CAPM, at about 1.5% per month, while the other factor models produce values closer to the average bill rate of 0.39% per month. This is robust to the choice of
lagged instruments. The zero beta rate is almost twice as large in the industry portfolio design than in the other portfolios.

In direct comparisons of factor models that do not use the test portfolios, we find that the FF3 factors significantly beat the market index and the FF5 model beats FF3, but the FF6 model with momentum only beats the FF5 when there is dynamic trading using a “modern” set of lagged instruments. The Q4 model beats FF5 in a fixed-weight comparison, but the two are not significantly different with dynamic trading. Dynamic trading improves the FF5 factors more than the Q4 factors. However, the Q5 model outperforms every model to which it is compared.
Appendix

A. Special Cases without Conditioning Information

Theorem I can be applied to cases with no conditioning information, such as fixed-weight factor models, with the following Corollary.

**Corollary I**: The asymptotic variance of the maximal estimated squared Sharpe Ratio $S_{\phi}^2$ with fixed weights $w$ and zero-beta rate $\phi$, with mean $\mu_{\phi}$, variance $\sigma_{\phi}^2$ and all estimated from the data, may be obtained using the Theorem I together with the following canonical matrices:

$$C = \frac{2(\mu_{\phi} - \phi)A^{-1}}{T\sigma_{\phi}^2} \sum_{t=1}^{T} Z_t^{-1} \left[ 1 - \frac{\mu_{\phi} - \phi}{\sigma_{\phi}^2} \left( \mu_{\phi} w - \mu_{\phi} \right) \right] w'$$

$$D = -\frac{(\mu_{\phi} - \phi)^2}{\sigma_{\phi}^4} w' = -\frac{S_{\phi}^2}{\sigma_{\phi}^2} w$$

One example of using Corollary I is testing the CAPM using, say, the value-weighted market portfolio. Corollary I delivers an asymptotic variance of the squared Sharpe ratio for the market portfolio. The weights vector $w$ is 1.0 on the market return and zero on the other test assets.

B. Details for Corollary II

The estimated squared Sharpe Ratio is

$$S_{\phi}^2 = \frac{\hat{\alpha}_2^2 + \hat{\alpha}_1 \hat{\alpha}_3 - 2\varphi \hat{\alpha}_2 + \varphi^2 (1 - \hat{\alpha}_3)}{\hat{\alpha}_1 (1 - \hat{\alpha}_3) - \hat{\alpha}_2^2}$$

which follows from maximizing
\[ \hat{S}^2 = \left( \hat{\mu}_p - \varphi \right)^2 \frac{\hat{\sigma}_p^2}{\hat{\sigma}_p^2} = \frac{\left( \hat{\mu}_p - \varphi \right)^2}{\hat{\alpha}_1 + \frac{\hat{\alpha}_2^2}{\hat{\alpha}_3} - 2\frac{\hat{\alpha}_2}{\hat{\alpha}_3} \hat{\mu}_p + \frac{1}{\hat{\alpha}_3} \hat{\mu}_p^2} \]

with respect to the mean \( \hat{\mu}_p \). The asymptotic variance of \( \hat{S}^2 \) follows from Theorem I together with the following choices for the canonical matrices \( C \) and \( D \):

\[
C = -\left[ \alpha_2 - \varphi (1 - \alpha_3) \right]^2 C_{\alpha_1} + 2 (\alpha_1 - \varphi \alpha_2) \left[ \alpha_2 - \varphi (1 - \alpha_3) \right] C_{\alpha_2} + (\alpha_1 - \varphi \alpha_2)^2 C_{\alpha_3} \\
D = -\left[ \alpha_2 - \varphi (1 - \alpha_3) \right]^2 D_{\alpha_1} + 2 (\alpha_1 - \varphi \alpha_2) \left[ \alpha_2 - \varphi (1 - \alpha_3) \right] D_{\alpha_2} + (\alpha_1 - \varphi \alpha_2)^2 D_{\alpha_3}
\]

C. Mimicking Portfolios and Cross-Sectional Fit

We use the same test assets whose cross-section we wish to explain to form the mimicking portfolios. One might think this will artificially improve the explanatory power. We show here it does not do so in the fixed-weight case.

In this section, we redefine \( f \) and \( r \) to be the \( T \times K \) and \( T \times N \) matrices of factor and test asset excess returns data. Represent the betas of the test assets on the factors as \( \beta = (f'f)^{-1} f'r \). (GLS works, too.) Mimicking portfolios are found by regressing the factors on the test assets. The mimicking excess returns are \( f^* = r (r'r)^{-1} r'f \). The betas of the test assets on the mimicking factor portfolios are

\[ \beta^* = (f^*f^*)^{-1}f^*r = [f'r (r'r)^{-1} r'f]^{-1} f'r. \]

Thus, the mimicking betas are related to the factor betas by an invertible \( K \times K \) rotation: \( \beta^* = A\beta \). A cross-sectional regression of returns on the factor betas delivers a coefficient \( \lambda' = (\beta\beta')^{-1}. \)

Thus, \( f'f' \) and fitted values \( \lambda \beta \). A cross-sectional regression of returns on the mimicking betas delivers a coefficient...
\[ \lambda^* = (\beta^* \beta^*)^{-1} \beta^* r^* = A^{-1} \lambda \] and fitted values \( \lambda^* \beta^* = \lambda \beta \).

With the same fitted values, there is no impact of using mimicking portfolios on the cross-sectional fit of the model.

If portfolios are formed to maximize the Sharpe ratio, there is an upward bias in any finite sample. This bias will artificially inflate the explanatory power, as measured by the fitted Sharpe ratio. One of our contributions, therefore, is bias adjustment for maximum Sharpe ratios. The mimicking portfolios with dynamic trading do not maximize the Sharpe ratio. We see no reason why they would artificially increase the explanatory power in the cross-section. Indeed, in Table 5 some of the mimicking portfolios display a negative finite sample bias.

**References**


Breeden, Douglas. “An intertemporal asset pricing model with stochastic consumption and


Table 1. Summary Statistics of Factors and Lagged Information
This table contains summary statistics for our sample of monthly traded factors from the French data library (February 1959 to December 2020), the Q factors from Hou, Xue, and Zhang (2015), and the nontraded factors. AR(1) is the first-order autocorrelation (after stochastic detrending when needed for the lagged instruments). Squared SR is the squared Sharpe ratio, where the zero-beta rate is the average Treasury bill rate, equal to 0.39 percent per month. The R-square is obtained by regressing market excess return on the lagged instruments. Returns, yields, and yield spreads are measured as percent per month.

<table>
<thead>
<tr>
<th>Traded Factors</th>
<th>Mean</th>
<th>Std</th>
<th>AR(1)</th>
<th>Squared SR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market-risk free</td>
<td>0.57</td>
<td>4.43</td>
<td>0.066</td>
<td>0.016</td>
</tr>
<tr>
<td>SMB</td>
<td>0.2</td>
<td>2.96</td>
<td>0.064</td>
<td>0.005</td>
</tr>
<tr>
<td>HML</td>
<td>0.26</td>
<td>2.81</td>
<td>0.179</td>
<td>0.008</td>
</tr>
<tr>
<td>RMW</td>
<td>0.23</td>
<td>2.09</td>
<td>0.149</td>
<td>0.012</td>
</tr>
<tr>
<td>CMA</td>
<td>0.24</td>
<td>1.92</td>
<td>0.121</td>
<td>0.015</td>
</tr>
<tr>
<td>Momentum</td>
<td>0.61</td>
<td>4.06</td>
<td>0.047</td>
<td>0.022</td>
</tr>
<tr>
<td>Investment</td>
<td>0.29</td>
<td>1.77</td>
<td>0.099</td>
<td>0.027</td>
</tr>
<tr>
<td>Profitability</td>
<td>0.44</td>
<td>2.4</td>
<td>0.117</td>
<td>0.034</td>
</tr>
<tr>
<td>Investment Growth</td>
<td>0.7</td>
<td>1.87</td>
<td>0.102</td>
<td>0.141</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Non-traded Factors</th>
<th>Mean</th>
<th>Std</th>
<th>AR(1)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Consumption growth</td>
<td>0.26</td>
<td>0.82</td>
<td>0.01</td>
<td>Na</td>
</tr>
<tr>
<td>Broker-dealer Leverage</td>
<td>0.09</td>
<td>6.71</td>
<td>0.09</td>
<td>Na</td>
</tr>
<tr>
<td>Expected Inflation</td>
<td>0.13</td>
<td>0.11</td>
<td>0.95</td>
<td>Na</td>
</tr>
<tr>
<td>Change Expected Inflation</td>
<td>0.00</td>
<td>0.03</td>
<td>0.19</td>
<td>Na</td>
</tr>
<tr>
<td>Industry Production</td>
<td>0.10</td>
<td>0.47</td>
<td>0.33</td>
<td>Na</td>
</tr>
<tr>
<td>Real Interest Rate</td>
<td>0.32</td>
<td>0.25</td>
<td>0.98</td>
<td>Na</td>
</tr>
<tr>
<td>Unexpected Inflation</td>
<td>0.00</td>
<td>0.12</td>
<td>0.18</td>
<td>Na</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Lagged Instruments</th>
<th>Mean</th>
<th>Std</th>
<th>AR(1)</th>
<th>R-Square (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short-Term T-bill Rate</td>
<td>-0.03</td>
<td>0.87</td>
<td>0.82</td>
<td>1.12</td>
</tr>
<tr>
<td>Investment Capital Ratio</td>
<td>0.00</td>
<td>0.00</td>
<td>0.86</td>
<td>0.05</td>
</tr>
<tr>
<td>Value Spread</td>
<td>0.00</td>
<td>0.08</td>
<td>0.73</td>
<td>0.18</td>
</tr>
<tr>
<td>Stock Return Dispersion</td>
<td>2.64</td>
<td>1.24</td>
<td>0.64</td>
<td>0.04</td>
</tr>
<tr>
<td>Net Equity Expansion</td>
<td>0.00</td>
<td>0.01</td>
<td>0.83</td>
<td>0.16</td>
</tr>
<tr>
<td>Relative Bill Rate</td>
<td>0.00</td>
<td>0.80</td>
<td>0.80</td>
<td>0.64</td>
</tr>
</tbody>
</table>
Table 2. Accuracy of Bias Adjustments for Squared Sharpe Ratios

The “true” squared Sharpe ratios are from simulations with a large number (1000*743) of time series observations. The values are stated in percent (multiplied by 100). The Average values across 5,000 simulation trials are shown for five alternative bias-adjustment methods. The number of time series observations in the finite samples is 743. The JK uses the results of Jobson and Korkie (1980), and are based on a Non-central F distribution. The four adjustments for dynamic portfolios are the Chi-square, Non-central F, Odd-even, and Direct Expansion. The adjustments are applied to the squared Sharpe ratios of fixed weight portfolios in Panel A and to Efficient with respect to Z portfolios in Panel B. The six lagged instruments that comprise the vector Z are described in the text and Table 1. The N=25 portfolios are the 5x5 size x book/market sorts, the N=49 are industry portfolios, and the N=99 combine the first two sets with 25 investment x profitability portfolios from Kenneth French.

Panel A: Fixed-weight Factor Portfolios

<table>
<thead>
<tr>
<th></th>
<th>TRUE</th>
<th>JK</th>
<th>% Diff</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rm</td>
<td>1.66</td>
<td>1.74</td>
<td>−5%</td>
</tr>
<tr>
<td>Sfx(FF3)</td>
<td>3.33</td>
<td>3.46</td>
<td>−4%</td>
</tr>
<tr>
<td>Sfx(FF5)</td>
<td>8.46</td>
<td>8.59</td>
<td>−2%</td>
</tr>
<tr>
<td>Sfx(FF6)</td>
<td>11.65</td>
<td>11.91</td>
<td>−2%</td>
</tr>
<tr>
<td>Sfx(Q5)</td>
<td>31.28</td>
<td>31.28</td>
<td>0%</td>
</tr>
<tr>
<td>Sfx(r)  N=25</td>
<td>16.35</td>
<td>16.55</td>
<td>−1%</td>
</tr>
<tr>
<td>Sfx(r) N=49</td>
<td>32.21</td>
<td>32.60</td>
<td>−1%</td>
</tr>
<tr>
<td>Sfx(r) N=99</td>
<td>77.19</td>
<td>78.23</td>
<td>−1%</td>
</tr>
</tbody>
</table>

Panel B: Efficient-with-Respect to Z Portfolios (% indicates differences from True)

<table>
<thead>
<tr>
<th></th>
<th>TRUE</th>
<th>No-Adj</th>
<th>Chi-Square</th>
<th>Non-Central F</th>
<th>Odd-Even</th>
<th>Direct Expansion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sux(FF3)</td>
<td>6.02</td>
<td>31.9%</td>
<td>24.3%</td>
<td>13.6%</td>
<td>12.6%</td>
<td>3.2%</td>
</tr>
<tr>
<td>Sux(FF5)</td>
<td>13.38</td>
<td>26.7%</td>
<td>20.4%</td>
<td>13.2%</td>
<td>11.9%</td>
<td>1.7%</td>
</tr>
<tr>
<td>Sux(FF6)</td>
<td>17.78</td>
<td>25.3%</td>
<td>19.3%</td>
<td>8.5%</td>
<td>7.1%</td>
<td>2.0%</td>
</tr>
<tr>
<td>Sux(Q5)</td>
<td>34.81</td>
<td>10.6%</td>
<td>7.6%</td>
<td>2.6%</td>
<td>1.6%</td>
<td>1.0%</td>
</tr>
<tr>
<td>Sux(r)  N=25</td>
<td>39.63</td>
<td>52.8%</td>
<td>38.2%</td>
<td>6.1%</td>
<td>0.4%</td>
<td>2.5%</td>
</tr>
<tr>
<td>Sux(r) N=49</td>
<td>90.2</td>
<td>47.1%</td>
<td>9.2%</td>
<td>−0.6%</td>
<td>8.7%</td>
<td>1.5%</td>
</tr>
<tr>
<td>Sux(r) N=99</td>
<td>184.07</td>
<td>58.9%</td>
<td>23.1%</td>
<td>1.5%</td>
<td>22.9%</td>
<td>2.7%</td>
</tr>
</tbody>
</table>
Table 3. Accuracy of Asymptotic Standard Deviations

A parametric bootstrap generates 1000 simulation trials, each with 743 observations. Squared Sharpe ratios and squared Sharpe ratio differences are estimated, and the asymptotic standard deviations are calculated using the propositions and Theorem I. The first columns (Empirical) are the standard deviations of the estimates across the 1,000 simulation trials. The Avg Asymptotic are the averages of the estimated asymptotic standard deviations. Fix(r) or fix(f) refers to a mean-variance efficient portfolio that ignores the conditioning information and uses fixed weights. UE is efficient with respect to Z. The lagged instruments are described in the data section. The average return of a three-month Treasury bill is taken to be the zero beta rate. The N=25 portfolios are the 5×5 size × book/market sorts, the N=49 are industry portfolios, and the N=99 combine the first two sets with 25 investment × profitability portfolios from Kenneth French.

Panel A: Standard Errors for Squared Sharpe Ratio Levels

<table>
<thead>
<tr>
<th></th>
<th>Empirical (simulated)</th>
<th>Average FSW Asymptotic</th>
<th>Average BKRS Asymptotic</th>
<th>Difference FSW (%) empirical</th>
<th>Difference BKRS (%) empirical</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_m$</td>
<td>0.28</td>
<td>0.28</td>
<td>0.28</td>
<td>−2%</td>
<td>−1%</td>
</tr>
<tr>
<td>$S_{fix}(FF3)$</td>
<td>0.45</td>
<td>0.40</td>
<td>0.41</td>
<td>−11%</td>
<td>−9%</td>
</tr>
<tr>
<td>$S_{fix}(FF6)$</td>
<td>0.87</td>
<td>0.80</td>
<td>0.81</td>
<td>−9%</td>
<td>−7%</td>
</tr>
<tr>
<td>$S_{UE}(FF3)$</td>
<td>0.60</td>
<td>0.58</td>
<td></td>
<td>−3%</td>
<td></td>
</tr>
<tr>
<td>$S_{UE}(FF6)$</td>
<td>1.09</td>
<td>1.08</td>
<td></td>
<td>−2%</td>
<td></td>
</tr>
<tr>
<td>$S_{fix}(r)$ N=25</td>
<td>1.02</td>
<td>0.95</td>
<td>0.95</td>
<td>−7%</td>
<td>−7%</td>
</tr>
<tr>
<td>$S_{fix}(r)$ N=49</td>
<td>1.25</td>
<td>1.24</td>
<td>1.30</td>
<td>0%</td>
<td>4%</td>
</tr>
<tr>
<td>$S_{fix}(r)$ N=99</td>
<td>2.33</td>
<td>2.48</td>
<td>2.54</td>
<td>7%</td>
<td>9%</td>
</tr>
<tr>
<td>$S_{UE}(r)$ N=25</td>
<td>1.49</td>
<td>1.38</td>
<td></td>
<td>−8%</td>
<td></td>
</tr>
<tr>
<td>$S_{UE}(r)$ N=49</td>
<td>1.91</td>
<td>1.79</td>
<td></td>
<td>−6%</td>
<td></td>
</tr>
<tr>
<td>$S_{UE}(r)$ N=99</td>
<td>3.50</td>
<td>3.33</td>
<td></td>
<td>−5%</td>
<td></td>
</tr>
</tbody>
</table>

Panel B: Standard Errors for Squared Sharpe Ratio Differences (N=25)

<table>
<thead>
<tr>
<th></th>
<th>Empirical (simulated)</th>
<th>Average FSW Asymptotic</th>
<th>Average BKRS Asymptotic</th>
<th>Difference FSW (%) empirical</th>
<th>Difference BKRS (%) empirical</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_{fix}(r) - R_m$</td>
<td>0.87</td>
<td>0.91</td>
<td>0.93</td>
<td>4%</td>
<td>6%</td>
</tr>
<tr>
<td>$S_{fix}(r) - S_{fix}(FF3)$</td>
<td>0.85</td>
<td>0.89</td>
<td>0.91</td>
<td>5%</td>
<td>7%</td>
</tr>
<tr>
<td>$S_{fix}(r) - S_{fix}(FF5)$</td>
<td>0.83</td>
<td>0.90</td>
<td>0.91</td>
<td>9%</td>
<td>10%</td>
</tr>
<tr>
<td>$S_{UE}(r) - R_m$</td>
<td>1.39</td>
<td>1.41</td>
<td></td>
<td>2%</td>
<td></td>
</tr>
<tr>
<td>$S_{UE}(r) - S_{UE}(FF3)$</td>
<td>1.41</td>
<td>1.40</td>
<td></td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>$S_{UE}(r) - S_{UE}(FF5)$</td>
<td>1.41</td>
<td>1.43</td>
<td></td>
<td>1%</td>
<td></td>
</tr>
<tr>
<td>$S_{UE}(r) - S_{fix}(r)$</td>
<td>1.06</td>
<td>1.04</td>
<td></td>
<td>−2%</td>
<td></td>
</tr>
</tbody>
</table>
Panel C: Standard Errors for Squared Sharpe Ratio Differences (N=99)

<table>
<thead>
<tr>
<th></th>
<th>Empirical (simulated)</th>
<th>Average Asymptotic</th>
<th>BKRS Asymptotic</th>
<th>Difference FSW (% empirical)</th>
<th>Difference BKRS (% empirical)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_{fix}(r) - R_m$</td>
<td>2.37</td>
<td>2.45</td>
<td>2.50</td>
<td>4%</td>
<td>6%</td>
</tr>
<tr>
<td>$S_{fix}(r) - S_{fix}(FF3)$</td>
<td>2.38</td>
<td>2.44</td>
<td>2.49</td>
<td>3%</td>
<td>5%</td>
</tr>
<tr>
<td>$S_{fix}(r) - S_{fix}(FF5)$</td>
<td>2.34</td>
<td>2.39</td>
<td>2.43</td>
<td>2%</td>
<td>4%</td>
</tr>
<tr>
<td>$S_{ue}(r) - R_m$</td>
<td>3.90</td>
<td>3.64</td>
<td></td>
<td>−7%</td>
<td></td>
</tr>
<tr>
<td>$S_{ue}(r) - S_{ue}(FF3)$</td>
<td>3.76</td>
<td>3.60</td>
<td></td>
<td>−4%</td>
<td></td>
</tr>
<tr>
<td>$S_{ue}(r) - S_{ue}(FF5)$</td>
<td>3.72</td>
<td>3.54</td>
<td></td>
<td>−5%</td>
<td></td>
</tr>
<tr>
<td>$S_{ue}(r) - S_{fix}(r)$</td>
<td>2.90</td>
<td>2.97</td>
<td></td>
<td>2%</td>
<td></td>
</tr>
</tbody>
</table>

Panel D: Standard Errors for Squared Sharpe Ratio Differences (Factors Alone)

<table>
<thead>
<tr>
<th></th>
<th>Empirical (simulated)</th>
<th>Average Asymptotic</th>
<th>BKRS Asymptotic</th>
<th>Difference FSW (% empirical)</th>
<th>Difference BKRS (% empirical)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_{fix}(FF5) - S_{fix}(FF3)$</td>
<td>0.54</td>
<td>0.48</td>
<td>0.49</td>
<td>−11%</td>
<td>−9%</td>
</tr>
<tr>
<td>$S_{fix}(FF6) - S_{fix}(FF5)$</td>
<td>0.52</td>
<td>0.47</td>
<td>0.47</td>
<td>−10%</td>
<td>−9%</td>
</tr>
<tr>
<td>$S_{fix}(FF6) - S_{fix}(Q5)$</td>
<td>1.06</td>
<td>1.02</td>
<td>1.03</td>
<td>−4%</td>
<td>−2%</td>
</tr>
<tr>
<td>$S_{ue}(FF5) - S_{ue}(FF3)$</td>
<td>0.63</td>
<td>0.68</td>
<td></td>
<td>8%</td>
<td></td>
</tr>
<tr>
<td>$S_{ue}(FF6) - S_{ue}(FF5)$</td>
<td>0.57</td>
<td>0.55</td>
<td></td>
<td>−3%</td>
<td></td>
</tr>
<tr>
<td>$S_{ue}(FF6) - S_{ue}(Q5)$</td>
<td>1.15</td>
<td>1.40</td>
<td></td>
<td>21%</td>
<td></td>
</tr>
</tbody>
</table>
Table 4. The Empirical Distributions of t-ratios

A parametric bootstrap generates 1000 simulation trials. Each set of simulated data has 743 observations. Squared Sharpe ratios, \( S(.) \), and their differences are estimated, bias adjusted using the second-order expansion method, and their asymptotic standard deviations are calculated. Squared t-ratios are formed as the squared adjusted Sharpe ratio or difference less its “true” value, divided by its asymptotic standard error. The true values are from simulations with 743*1000 observations. Critical values of the empirical distribution from the 1,000 simulation trials are shown. \( \chi(1) \) are the values for a Chi distribution with one degree of freedom. Fixed weight portfolios ignore the conditioning information. UE is efficient with respect to \( Z \). The lagged instruments are described in the text. The average return of a three-month Treasury bill is taken to be the zero beta rate. The N=25 portfolios are the 5×5 size × book/market sorts, the N=49 are industry portfolios and the N=99 combine the first two sets with 25 investment × profitability portfolios from Kenneth French.

Panel A: Critical Values for T-ratios of Squared Sharpe Ratio Levels

<table>
<thead>
<tr>
<th>Percentile: ( \chi(1) )</th>
<th>Fixed Weight Portfolios</th>
<th>Dynamic UE Portfolios</th>
</tr>
</thead>
<tbody>
<tr>
<td>90%</td>
<td>95%</td>
<td>98%</td>
</tr>
<tr>
<td>1.65</td>
<td>1.96</td>
<td>2.33</td>
</tr>
<tr>
<td>2.23</td>
<td>2.96</td>
<td>3.70</td>
</tr>
<tr>
<td>1.76</td>
<td>2.22</td>
<td>2.72</td>
</tr>
<tr>
<td>1.90</td>
<td>2.21</td>
<td>2.79</td>
</tr>
<tr>
<td>1.85</td>
<td>2.24</td>
<td>2.50</td>
</tr>
<tr>
<td>1.78</td>
<td>2.16</td>
<td>2.58</td>
</tr>
</tbody>
</table>

| \( S(r) \) N=25 | 1.62 | 1.97 | 2.21 | 1.62 | 1.97 | 2.21 |
| \( S(r) \) N=49 | 1.59 | 1.88 | 2.28 | 1.82 | 2.12 | 2.49 |
| \( S(r) \) N=99 | 1.65 | 1.95 | 2.24 | 1.85 | 2.18 | 2.56 |

Panel B: Squared Sharpe Ratio Differences (N=25)

<table>
<thead>
<tr>
<th>Percentile: ( \chi(1) )</th>
<th>Fixed Weight Portfolios</th>
<th>Dynamic UE Portfolios</th>
</tr>
</thead>
<tbody>
<tr>
<td>90%</td>
<td>95%</td>
<td>98%</td>
</tr>
<tr>
<td>1.65</td>
<td>1.96</td>
<td>2.33</td>
</tr>
<tr>
<td>1.55</td>
<td>1.88</td>
<td>2.27</td>
</tr>
<tr>
<td>1.56</td>
<td>1.86</td>
<td>2.21</td>
</tr>
<tr>
<td>1.51</td>
<td>1.84</td>
<td>2.27</td>
</tr>
<tr>
<td>SUE(r) − Sfix(r)</td>
<td>1.62</td>
<td>1.87</td>
</tr>
</tbody>
</table>
Panel C: Squared Sharpe Ratio Differences (N=49)

<table>
<thead>
<tr>
<th>Percentile:</th>
<th>Fixed Weight Portfolios</th>
<th>Dynamic UE Portfolios</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>90%  95%  98%</td>
<td>90%  95%  98%</td>
</tr>
<tr>
<td>$\chi(1)$</td>
<td>1.65  1.96  2.33</td>
<td>1.65  1.96  2.33</td>
</tr>
<tr>
<td>S(r) – Rm</td>
<td>1.57  1.91  2.27</td>
<td>1.77  2.15  2.52</td>
</tr>
<tr>
<td>S(r) – S(FF3)</td>
<td>1.58  1.91  2.13</td>
<td>1.85  2.16  2.51</td>
</tr>
<tr>
<td>S(r) – S(FF5)</td>
<td>1.57  1.90  2.22</td>
<td>1.82  2.16  2.58</td>
</tr>
<tr>
<td>SUE(r) – Sfix(r)</td>
<td>1.72  2.17  2.53</td>
<td>NA     NA     NA</td>
</tr>
</tbody>
</table>

Panel D: Squared Sharpe Ratio Differences (N=99)

<table>
<thead>
<tr>
<th>Percentile:</th>
<th>Fixed Weight Portfolios</th>
<th>Dynamic UE Portfolios</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>90%  95%  98%</td>
<td>90%  95%  98%</td>
</tr>
<tr>
<td>$\chi(1)$</td>
<td>1.65  1.96  2.33</td>
<td>1.65  1.96  2.33</td>
</tr>
<tr>
<td>S(r) – Rm</td>
<td>2.04  2.34  2.67</td>
<td>1.86  2.23  2.50</td>
</tr>
<tr>
<td>S(r) – S(FF3)</td>
<td>1.68  1.98  2.32</td>
<td>1.84  2.14  2.50</td>
</tr>
<tr>
<td>S(r) – S(FF5)</td>
<td>1.77  2.10  2.54</td>
<td>1.77  2.19  2.54</td>
</tr>
<tr>
<td>SUE(r) – Sfix(r)</td>
<td>1.65  2.01  2.51</td>
<td>NA     NA     NA</td>
</tr>
</tbody>
</table>

Panel E: Squared Sharpe Ratio Differences for Factors Alone

<table>
<thead>
<tr>
<th>Percentile:</th>
<th>Fixed Weight Portfolios</th>
<th>Dynamic UE Portfolios</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>90%  95%  98%</td>
<td>90%  95%  98%</td>
</tr>
<tr>
<td>$\chi(1)$</td>
<td>1.65  1.96  2.33</td>
<td>1.65  1.96  2.33</td>
</tr>
<tr>
<td>S(FF5) – S(FF3)</td>
<td>1.95  2.38  2.96</td>
<td>1.54  1.83  2.42</td>
</tr>
<tr>
<td>S(FF6) – S(FF5)</td>
<td>2.08  2.65  3.45</td>
<td>1.84  2.22  2.86</td>
</tr>
<tr>
<td>S(FF6) – S(Q5)</td>
<td>1.70  1.98  2.50</td>
<td>1.39  1.73  2.08</td>
</tr>
</tbody>
</table>
### Table 5. Tests of Factor Model Efficiency

The test statistic is the difference in the bias-adjusted squared Sharpe ratios (not multiplied by 100) for the test assets and the factors versus the factors alone. The t-ratios are in parentheses. The factor model abbreviations and test asset portfolios are described in the text. Monthly Sharpe ratios are computed using the average Treasury bill return of 0.39 percent per month as the zero-beta rate. The dynamic models trade optimally using the lagged instruments. The sample period is January 1967 to December, 2013.

<table>
<thead>
<tr>
<th>Panel A: Fixed weight Models and Assets</th>
<th>CAPM</th>
<th>FF3</th>
<th>FF5</th>
<th>FF6</th>
<th>Q4</th>
<th>Q5</th>
<th>CRR3</th>
</tr>
</thead>
<tbody>
<tr>
<td>25 size x value portfolios:</td>
<td>0.15</td>
<td>0.12</td>
<td>0.10</td>
<td>0.08</td>
<td>0.11</td>
<td>0.10</td>
<td>0.12</td>
</tr>
<tr>
<td>(3.92)</td>
<td>(3.32)</td>
<td>(2.80)</td>
<td>(2.45)</td>
<td>(2.99)</td>
<td>(2.53)</td>
<td>(3.45)</td>
<td></td>
</tr>
<tr>
<td>25 Investment x profitability:</td>
<td>0.07</td>
<td>0.05</td>
<td>0.00</td>
<td>−0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.05</td>
</tr>
<tr>
<td>(2.34)</td>
<td>(1.87)</td>
<td>(−0.20)</td>
<td>(−0.67)</td>
<td>(0.34)</td>
<td>(0.38)</td>
<td>(1.43)</td>
<td></td>
</tr>
<tr>
<td>32 size x value portfolios x prod:</td>
<td>0.10</td>
<td>0.07</td>
<td>0.05</td>
<td>0.05</td>
<td>0.06</td>
<td>0.06</td>
<td>0.07</td>
</tr>
<tr>
<td>(2.85)</td>
<td>(2.27)</td>
<td>(1.55)</td>
<td>(1.45)</td>
<td>(1.82)</td>
<td>(1.56)</td>
<td>(2.32)</td>
<td></td>
</tr>
<tr>
<td>49 Industry:</td>
<td>0.08</td>
<td>0.12</td>
<td>0.17</td>
<td>0.15</td>
<td>0.14</td>
<td>0.12</td>
<td>0.04</td>
</tr>
<tr>
<td>(2.48)</td>
<td>(3.07)</td>
<td>(3.78)</td>
<td>(3.37)</td>
<td>(2.88)</td>
<td>(2.52)</td>
<td>(1.28)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Dynamic Models vs Dynamic Assets</th>
<th>CAPM</th>
<th>FF3</th>
<th>FF5</th>
<th>FF6</th>
<th>Q4</th>
<th>Q5</th>
<th>CRR3</th>
</tr>
</thead>
<tbody>
<tr>
<td>25 size x value portfolios:</td>
<td>0.31</td>
<td>0.28</td>
<td>0.28</td>
<td>0.27</td>
<td>0.29</td>
<td>0.27</td>
<td>0.22</td>
</tr>
<tr>
<td>(5.11)</td>
<td>(4.67)</td>
<td>(4.66)</td>
<td>(4.61)</td>
<td>(4.60)</td>
<td>(4.03)</td>
<td>(3.89)</td>
<td></td>
</tr>
<tr>
<td>25 Investment x profitability:</td>
<td>0.19</td>
<td>0.18</td>
<td>0.12</td>
<td>0.11</td>
<td>0.13</td>
<td>0.14</td>
<td>0.09</td>
</tr>
<tr>
<td>(3.64)</td>
<td>(3.39)</td>
<td>(2.75)</td>
<td>(2.67)</td>
<td>(2.91)</td>
<td>(2.60)</td>
<td>(1.73)</td>
<td></td>
</tr>
<tr>
<td>32 size x value portfolios x prod:</td>
<td>0.22</td>
<td>0.18</td>
<td>0.15</td>
<td>0.15</td>
<td>0.19</td>
<td>0.19</td>
<td>0.15</td>
</tr>
<tr>
<td>(3.48)</td>
<td>(3.03)</td>
<td>(2.98)</td>
<td>(3.00)</td>
<td>(3.20)</td>
<td>(2.90)</td>
<td>(2.60)</td>
<td></td>
</tr>
<tr>
<td>49 Industry:</td>
<td>0.23</td>
<td>0.26</td>
<td>0.33</td>
<td>0.31</td>
<td>0.30</td>
<td>0.29</td>
<td>0.17</td>
</tr>
<tr>
<td>(3.86)</td>
<td>(4.15)</td>
<td>(4.63)</td>
<td>(4.52)</td>
<td>(4.14)</td>
<td>(3.90)</td>
<td>(2.89)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel C: Dynamic Assets vs Full Conditional Models</th>
<th>CAPM</th>
<th>FF3</th>
<th>FF5</th>
<th>FF6</th>
<th>Q4</th>
<th>Q5</th>
<th>CRR3</th>
</tr>
</thead>
<tbody>
<tr>
<td>25 size x value portfolios:</td>
<td>0.24</td>
<td>0.24</td>
<td>0.24</td>
<td>0.24</td>
<td>0.24</td>
<td>0.24</td>
<td>0.25</td>
</tr>
<tr>
<td>25 Investment x profitability:</td>
<td>0.11</td>
<td>0.11</td>
<td>0.11</td>
<td>0.11</td>
<td>0.11</td>
<td>0.11</td>
<td>0.13</td>
</tr>
<tr>
<td>(2.17)</td>
<td>(2.17)</td>
<td>(2.17)</td>
<td>(2.17)</td>
<td>(2.18)</td>
<td>(2.18)</td>
<td>(2.49)</td>
<td></td>
</tr>
<tr>
<td>32 size x value portfolios x prod:</td>
<td>0.19</td>
<td>0.19</td>
<td>0.19</td>
<td>0.19</td>
<td>0.19</td>
<td>0.19</td>
<td>0.18</td>
</tr>
<tr>
<td>(3.16)</td>
<td>(3.16)</td>
<td>(3.16)</td>
<td>(3.16)</td>
<td>(3.15)</td>
<td>(3.15)</td>
<td>(3.08)</td>
<td></td>
</tr>
<tr>
<td>49 Industry:</td>
<td>0.19</td>
<td>0.19</td>
<td>0.19</td>
<td>0.19</td>
<td>0.19</td>
<td>0.19</td>
<td>0.20</td>
</tr>
<tr>
<td>(3.43)</td>
<td>(3.43)</td>
<td>(3.43)</td>
<td>(3.43)</td>
<td>(3.43)</td>
<td>(3.43)</td>
<td>(3.49)</td>
<td></td>
</tr>
</tbody>
</table>
Table 6. Tests of Factor Models with an Estimated Zero Beta Rate

The test statistic is the difference in bias-adjusted squared Sharpe ratios for the test assets and the factors versus the factors alone. The factor model abbreviations and test asset portfolios are described in the text. Monthly Sharpe ratios are computed using the estimated zero beta rate. The dynamic models trade optimally using the lagged instruments following Cooper and Maio (2019), January 1967 to December, 2013.

### Panel A: Fixed weight Models

<table>
<thead>
<tr>
<th></th>
<th>CAPM</th>
<th>FF3</th>
<th>FF5</th>
<th>FF6</th>
<th>Q4</th>
<th>Q5</th>
<th>CRR3</th>
</tr>
</thead>
<tbody>
<tr>
<td>25 size x value portfolios:</td>
<td>0.05</td>
<td>0.06</td>
<td>0.06</td>
<td>0.05</td>
<td>0.06</td>
<td>0.08</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>(1.94)</td>
<td>(1.00)</td>
<td>(0.69)</td>
<td>(0.49)</td>
<td>(0.95)</td>
<td>(0.73)</td>
<td>(0.63)</td>
</tr>
<tr>
<td>Zero Beta rate</td>
<td>0.0152</td>
<td>0.0040</td>
<td>0.0039</td>
<td>0.0038</td>
<td>0.0032</td>
<td>0.0028</td>
<td>0.0076</td>
</tr>
<tr>
<td>25 Investment x profitability:</td>
<td>0.03</td>
<td>0.04</td>
<td>-0.01</td>
<td>-0.02</td>
<td>-0.01</td>
<td>-0.01</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(1.18)</td>
<td>(1.67)</td>
<td>(-0.15)</td>
<td>(-0.31)</td>
<td>(-0.12)</td>
<td>(-0.14)</td>
<td>(0.12)</td>
</tr>
<tr>
<td>Zero Beta rate</td>
<td>0.0156</td>
<td>0.0054</td>
<td>0.0041</td>
<td>0.0039</td>
<td>0.0048</td>
<td>0.0047</td>
<td>0.0070</td>
</tr>
<tr>
<td>32 size x value portfolios x prod:</td>
<td>0.03</td>
<td>0.04</td>
<td>0.02</td>
<td>0.02</td>
<td>0.03</td>
<td>0.04</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(1.39)</td>
<td>(0.66)</td>
<td>(0.26)</td>
<td>(0.21)</td>
<td>(0.55)</td>
<td>(0.39)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>Zero Beta rate</td>
<td>0.0116</td>
<td>0.0044</td>
<td>0.0041</td>
<td>0.0039</td>
<td>0.0039</td>
<td>0.0036</td>
<td>0.0054</td>
</tr>
<tr>
<td>49 Industry:</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.09</td>
<td>0.08</td>
<td>0.09</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>(2.51)</td>
<td>(2.27)</td>
<td>(1.73)</td>
<td>(1.75)</td>
<td>(2.14)</td>
<td>(1.92)</td>
<td>(0.35)</td>
</tr>
<tr>
<td>Zero Beta rate</td>
<td>0.0083</td>
<td>0.0074</td>
<td>0.0066</td>
<td>0.0061</td>
<td>0.0080</td>
<td>0.0072</td>
<td>0.0050</td>
</tr>
</tbody>
</table>

### Panel B: Dynamic Models

<table>
<thead>
<tr>
<th></th>
<th>CAPM</th>
<th>FF3</th>
<th>FF5</th>
<th>FF6</th>
<th>Q4</th>
<th>Q5</th>
<th>CRR3</th>
</tr>
</thead>
<tbody>
<tr>
<td>25 size x value portfolios:</td>
<td>0.16</td>
<td>0.17</td>
<td>0.16</td>
<td>0.15</td>
<td>0.18</td>
<td>0.19</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>(2.96)</td>
<td>(2.27)</td>
<td>(1.67)</td>
<td>(1.46)</td>
<td>(2.23)</td>
<td>(1.67)</td>
<td>(1.82)</td>
</tr>
<tr>
<td>Zero Beta rate</td>
<td>0.0147</td>
<td>0.0041</td>
<td>0.0039</td>
<td>0.0038</td>
<td>0.0032</td>
<td>0.0028</td>
<td>0.0074</td>
</tr>
<tr>
<td>25 Investment x profitability:</td>
<td>0.08</td>
<td>0.10</td>
<td>0.04</td>
<td>0.03</td>
<td>0.04</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>(1.76)</td>
<td>(2.08)</td>
<td>(0.61)</td>
<td>(0.40)</td>
<td>(0.70)</td>
<td>(0.42)</td>
<td>(0.62)</td>
</tr>
<tr>
<td>Zero Beta rate</td>
<td>0.0155</td>
<td>0.0054</td>
<td>0.0040</td>
<td>0.0039</td>
<td>0.0047</td>
<td>0.0046</td>
<td>0.0069</td>
</tr>
<tr>
<td>32 size x value portfolios x prod:</td>
<td>0.12</td>
<td>0.10</td>
<td>0.09</td>
<td>0.08</td>
<td>0.12</td>
<td>0.12</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>(2.17)</td>
<td>(1.55)</td>
<td>(0.90)</td>
<td>(0.84)</td>
<td>(1.68)</td>
<td>(1.22)</td>
<td>(1.13)</td>
</tr>
<tr>
<td>Zero Beta rate</td>
<td>0.0114</td>
<td>0.0044</td>
<td>0.0041</td>
<td>0.0040</td>
<td>0.0038</td>
<td>0.0036</td>
<td>0.0053</td>
</tr>
<tr>
<td>49 Industry:</td>
<td>0.19</td>
<td>0.18</td>
<td>0.19</td>
<td>0.21</td>
<td>0.20</td>
<td>0.21</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>(3.25)</td>
<td>(3.23)</td>
<td>(3.16)</td>
<td>(3.34)</td>
<td>(3.59)</td>
<td>(3.52)</td>
<td>(1.89)</td>
</tr>
<tr>
<td>Zero Beta rate</td>
<td>0.0081</td>
<td>0.0074</td>
<td>0.0066</td>
<td>0.0061</td>
<td>0.0080</td>
<td>0.0071</td>
<td>0.0049</td>
</tr>
</tbody>
</table>
Table 7. Direct Factor Model Comparisons
The test statistic is the difference in bias-adjusted squared Sharpe ratios (not multiplied by 100) for the first model less the second model. The factors are held with fixed weights over time (no instruments) or dynamically traded using the various sets of instruments. The t-ratios in parentheses are the differences divided by the asymptotic standard errors for the difference. The factor model abbreviations and test asset portfolios are described in the text. FF6* replaces the HML factor in FF6 with a monthly rebalanced version. The sample period is January 1967 to December, 2020.

<table>
<thead>
<tr>
<th></th>
<th>No Instruments</th>
<th>Classical Instruments</th>
<th>Modern Instruments</th>
<th>Cooper Miao</th>
</tr>
</thead>
<tbody>
<tr>
<td>FF3-Rm</td>
<td>0.03</td>
<td>0.05</td>
<td>0.06</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>(1.73)</td>
<td>(2.45)</td>
<td>(2.74)</td>
<td>(1.96)</td>
</tr>
<tr>
<td>FF5-FF3</td>
<td>0.06</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>(2.38)</td>
<td>(2.74)</td>
<td>(2.44)</td>
<td>(2.04)</td>
</tr>
<tr>
<td>FF6-FF5</td>
<td>0.03</td>
<td>0.03</td>
<td>0.06</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>(1.37)</td>
<td>(1.46)</td>
<td>(2.40)</td>
<td>(1.54)</td>
</tr>
<tr>
<td>Q5-Q4</td>
<td>0.20</td>
<td>0.20</td>
<td>0.21</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>(3.75)</td>
<td>(3.80)</td>
<td>(3.81)</td>
<td>(3.52)</td>
</tr>
<tr>
<td>Q4-FF5</td>
<td>0.08</td>
<td>0.03</td>
<td>0.06</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>(2.20)</td>
<td>(0.81)</td>
<td>(1.36)</td>
<td>(1.65)</td>
</tr>
<tr>
<td>Q5-FF6</td>
<td>0.25</td>
<td>0.21</td>
<td>0.20</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>(4.05)</td>
<td>(3.23)</td>
<td>(3.09)</td>
<td>(3.46)</td>
</tr>
<tr>
<td>Q5-FF6*</td>
<td>0.22</td>
<td>0.19</td>
<td>0.19</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>(3.27)</td>
<td>(2.86)</td>
<td>(2.81)</td>
<td>(2.84)</td>
</tr>
<tr>
<td>Q5-CCR3</td>
<td>0.31</td>
<td>0.26</td>
<td>0.32</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td>(3.71)</td>
<td>(3.07)</td>
<td>(3.37)</td>
<td>(3.40)</td>
</tr>
</tbody>
</table>
Internet Appendix for:

Factor Model Comparisons with Conditioning Information

April 9, 2023

This Internet Appendix provides proofs and ancilliary results for the named paper. Section I provides detailed descriptions of bias corrections and proofs of the main propositions. Section 2 derives results for the case with time-varying covariances. Section 3 establishes the equivalence between two formulations of the efficiency with respect to Z portfolio optimization. Section 4 presents a decomposition of the Sharpe ratio improvements with dynamic trading. Section 5 presents simulations and empirical results using alternative choices of lagged instruments. Section 6 presents results for mimicking portfolios for nontraded factors.

1. Proofs of the Main results

Plug-in Bias Adjustments

The first bias adjustment is based on the noncentral Chi-square distribution, assuming normality and a given covariance matrix. The estimated conditional mean return is \( \hat{\delta} ' Z_t \) and its conditional variance is \( Q = \text{Var}( (\hat{\delta} - \delta)' Z_t | Z) \). The quadratic form \( ((\hat{\delta} - \delta)' Z_t)' Q^{-1} ((\hat{\delta} - \delta)' Z_t) \) is central Chi-square with N degrees of freedom. The quadratic form \( ((\hat{\delta} - \delta)' Z_t)' V^{-1} ((\hat{\delta} - \delta)' Z_t) \) has mean equal to \( [Z_t (Z'Z)^{-1} Z_t] \) and the sample squared conditional Sharpe ratio, \( \hat{\delta} ' Z_t)' V^{-1} \hat{\delta} ' Z_t \) has mean \( [\hat{\delta}'Z_t]' V^{-1} [\hat{\delta}'Z_t] + N [Z_t (Z'Z)^{-1} Z_t] \). The bias-adjusted conditional ratio is:

\[
S_{1*2}^2(Z_t) = [ \hat{\delta} ' Z_t]' V^{-1} [\hat{\delta} ' Z_t] - N [Z_t (Z'Z)^{-1} Z_t],
\]

where the true variance matrix \( V \) is assumed in this adjustment. The adjusted UE maximum squared Sharpe ratio is then found as \( \frac{1}{T} \sum_{t=1}^{T} (1 + S_{1*2}^2(Z_t))^{-1} -1 \).

The second bias adjustment works similarly, but following Jobson and Korkie (1980) uses the assumption of a non-central \( F \) distribution, accommodating estimation error in the covariance matrix. If the estimated conditional mean return \( \hat{\delta} Z_t \) is normal and the estimated conditional variance is \( \hat{V} \), then...
where $\sum_{t=1}^{T} \hat{e}_t \hat{e}'_t$ has a Wishart distribution with scale matrix $V$ along with degrees of freedom $T - L$. The sample squared conditional Sharpe ratio has expectation

$$E \left( \hat{V}^{-1} \right) = TE \left( \frac{\sum_{t=1}^{T} \hat{e}_t \hat{e}'_t}{T - N - L - 1} \right) = \frac{T}{T - N - L} \hat{V}^{-1} \hat{V} \equiv \frac{1}{T} \sum_{t=1}^{T} \left( R_t - \hat{\delta}'Z_t \right) \left( R_t - \hat{\delta}'Z_t \right)' = \frac{1}{T} \sum_{t=1}^{T} \hat{e}_t \hat{e}'_t$$

Solving for an unbiased estimator we obtain our second bias adjusted estimator for the conditional ratio:

$$S_{2*}^2(Z_t) = \left( \hat{\delta}'Z_t \right)' \hat{V}^{-1} \left( \hat{\delta}'Z_t \right) \frac{T - N - L - 1}{T} - N \left[ Z_t (Z'Z)^{-1} Z_t \right]$$

As before, the adjusted UE maximum squared Sharpe ratio is found as $\left[ \frac{1}{T} \sum_{t=1}^{T} \left( 1 + S_{2*}^2(Z_t) \right)^{-1} \right]^{-1}$. Note that even if an unbiased conditional Sharpe ratio is used in the first two adjustments, the UE Sharpe ratio can retain some bias because of its nonlinear relation to the conditional ratios.

The third bias adjustment follows Noh et al. (2019) by using odd versus even months to estimate the regression coefficients. The bias of $\left[ \hat{\delta}'Z_t \right]' \hat{V}^{-1} \left[ \hat{\delta}'Z_t \right]$ comes in part from nonlinearity, via Jensens’ inequality, and partly from the covariance between estimation errors in the two $\left[ \hat{\delta}'Z_t \right]$. Our third bias adjustment method splits the time-series into subsamples with odd months and even months, and estimates the two terms using the subsamples. Then

$$S_{3*}^2(Z_t) = \left[ \hat{\delta}_{\text{odd}}'Z_t \right]' \hat{V}^{-1} \left[ \hat{\delta}_{\text{even}}'Z_t \right].$$

If the autocorrelation of returns are fully captured by $\hat{\delta}'Z_t$, and residuals in regression (10) are not autocorrelated, the errors in odd and even months should be uncorrelated, which should reduce the finite sample bias. This approach, however may lose some efficiency.

**B. Partial derivatives for Proposition II:**

$$\frac{\partial S_{\phi}^2}{\partial \alpha_1} = -\left( \frac{\alpha_2 - \phi \alpha_3}{\alpha_1 \alpha_3 - \alpha_2^2} \right)^2, \quad \frac{\partial S_{\phi}^2}{\partial \alpha_2} = 2 \left( \frac{\alpha_1 - \phi \alpha_2}{\alpha_1 \alpha_3 - \alpha_2^2} \right) \left( \alpha_2 - \phi \alpha_3 \right), \quad \frac{\partial S_{\phi}^2}{\partial \alpha_3} = -\left( \frac{\alpha_3 - \phi \alpha_2}{\alpha_1 \alpha_3 - \alpha_2^2} \right)^2,$$
\[
\frac{\partial^2 S^2_\phi}{\partial \alpha_i^2} = 2 \alpha_i \left( \alpha_i^2 - \varphi \alpha_i \right)^2, \quad \frac{\partial^2 S^2_\phi}{\partial \alpha_2^2} = 2 \alpha_2 \left( \alpha_2 - \varphi \alpha_2 \right)^2 + 8 \alpha_2 \left( \alpha_1 - \varphi \alpha_2 \right) \left( \alpha_2 - \varphi \alpha_3 \right),
\]
\[
\frac{\partial^2 S^2_\phi}{\partial \alpha_3^2} = 2 \alpha_3 \left( \alpha_3 - \varphi \alpha_3 \right)^2, \quad \frac{\partial^2 S^2_\phi}{\partial \alpha_i \partial \alpha_2} = -2 \left( \alpha_2 - \varphi \alpha_3 \right) \alpha_i \left( \alpha_1 - 2 \varphi \alpha_2 \right),
\]
\[
\frac{\partial^2 S^2_\phi}{\partial \alpha_i \partial \alpha_3} = 2 \alpha_i \left( \alpha_i - \varphi \alpha_i \right) \left( \alpha_2 - \varphi \alpha_3 \right) \text{ and } \frac{\partial^2 S^2_\phi}{\partial \alpha_2 \partial \alpha_3} = 2 \frac{\alpha_2 - \varphi \alpha_2}{\alpha_1 \left( \alpha_1 - \alpha_2 \right)^2} \left( \varphi - 2 \frac{\alpha_2 \left( \alpha_1 - \varphi \alpha_2 \right)}{\alpha_1 \alpha_3 - \alpha_2^2} \right).
\]

Where \( a = 1'V^{-1} \mu, \ b_i = 1'V^{-1} \mu_i, \ c_k = 1 + \mu_k'V^{-1} \mu_i, \) and \( d_{ik} = ac_{ik} - b_k b_i \) and their estimates are \( \hat{a} = 1'\hat{V}^{-1} \mu, \ \hat{b}_i = 1'\hat{V}^{-1} \mu_i, \ \hat{c}_k = 1 + \hat{\mu}_k'\hat{V}^{-1} \hat{\mu}_i, \) where \( \hat{d}_{ik} = \hat{a}\hat{c}_{ik} - \hat{b}_k \hat{b}_i. \) Then the consistent estimates are
\[
\hat{\alpha}_1 = \frac{1}{T} \sum_{t=1}^{T} \hat{c}_n, \quad \hat{\alpha}_2 = \frac{1}{T} \sum_{t=1}^{T} \hat{b}_n, \quad \text{and} \quad \hat{\alpha}_3 = \frac{1}{T} \sum_{t=1}^{T} \frac{1}{d_n}.
\]

**Applications of Asymptotic Variances for Optimal Zero-beta rates**

**Case 1: Test for Squared Sharpe ratios of two UE portfolios.**

The true squared Sharpe ratio for a UE portfolio is
\[
S^2_\phi = \frac{\alpha_1^2 + \alpha_2 - \varphi \alpha_2 + \varphi^2 \left( 1 - \alpha_3 \right)}{\alpha_1 \left( 1 - \alpha_3 - \alpha_2^2 \right)}. \tag{A.4}
\]

The squared Sharpe ratio is a quadratic function of \( \phi. \) Hence, we obtain:
\[
a = \frac{\alpha_1^2 + \alpha_2 - \varphi \alpha_2}{\alpha_1 \left( 1 - \alpha_3 - \alpha_2^2 \right)}, \tag{A.5}
\]
\[
b = \frac{\alpha_2}{\alpha_1 \left( 1 - \alpha_3 - \alpha_2^2 \right)}, \tag{A.6}
\]
\[
c = \frac{1 - \alpha_3}{\alpha_1 \left( 1 - \alpha_3 - \alpha_2^2 \right)}. \tag{A.7}
\]

Since \( a, b, \) and \( c \) are all continuous differentiable functions of \( \alpha_1, \alpha_2, \) and \( \alpha_3, \) we can expand \( \hat{a}, \hat{a}^*, \hat{b}, \hat{b}^*, \hat{c} \) and \( \hat{c}^*: \)
\[
\hat{a} - a = -\frac{\alpha_2^2}{\alpha_1 \left( 1 - \alpha_3 - \alpha_2^2 \right)} \left( \hat{\alpha}_1 - \alpha_1 \right) +
\]
\[
\frac{2\alpha_1 \alpha_2}{\alpha_1 \left( 1 - \alpha_3 - \alpha_2^2 \right)} \left( \hat{\alpha}_2 - \alpha_2 \right) + \frac{\alpha_1^2}{\alpha_1 \left( 1 - \alpha_3 - \alpha_2^2 \right)} \left( \hat{\alpha}_3 - \alpha_3 \right) + O \left( \frac{1}{T} \right),
\]
\[
\hat{a}^* - a^* = -\frac{(a_2^*)^*}{(a_1(1-a_3) - (a_2^*)^2)} (\hat{a}_4^* - a_1^*) + \\
\frac{2\alpha_1\alpha_2^2}{(a_1(1-a_3) - (a_2^*)^2)} (\hat{a}_2^* - a_2^*) + \frac{(a_1^*)^2}{(a_1(1-a_3) - (a_2^*)^2)} (\hat{a}_3^* - a_3^*) + O\left(\frac{1}{T}\right),
\]
(A.8)

\[
\hat{b} - b = -\frac{\alpha_2(1-a_3)}{(a_1(1-a_3) - a_2^2)} (\hat{a}_1 - a_1) + \\
\frac{\alpha_1(1-a_3)+a_2^2}{(a_1(1-a_3) - a_2^2)} (\hat{a}_2 - a_2) + \frac{\alpha_3\alpha_2}{(a_1(1-a_3) - a_2^2)} (\hat{a}_3 - a_3) + O\left(\frac{1}{T}\right),
\]
(A.9)

\[
\hat{b}^* - b^* = -\frac{\alpha_2(1-a_3)}{(a_1(1-a_3) - a_2^2)} (\hat{a}_1^* - a_1^*) + \\
\frac{\alpha_1(1-a_3)+a_2^2}{(a_1(1-a_3) - a_2^2)} (\hat{a}_2^* - a_2^*) + \frac{\alpha_3\alpha_2}{(a_1(1-a_3) - a_2^2)} (\hat{a}_3^* - a_3^*) + O\left(\frac{1}{T}\right),
\]
(A.10)

\[
\hat{c} - c = \frac{-\alpha_3^2}{(a_1(1-a_3) - a_2^2)} (\hat{a}_4 - a_1) + \frac{2\alpha_2(1-a_3)}{(a_1(1-a_3) - a_2^2)} (\hat{a}_2 - a_2) \\
+ \frac{\alpha_3^2}{(a_1(1-a_3) - a_2^2)} (\hat{a}_3 - a_3) + O\left(\frac{1}{T}\right),
\]
(A.11)

\[
\hat{c}^* - c^* = \frac{-\alpha_3^2}{(a_1(1-a_3) - a_2^2)} (\hat{a}_4^* - a_1^*) + \frac{2\alpha_2(1-a_3)}{(a_1(1-a_3) - a_2^2)} (\hat{a}_2^* - a_2^*) \\
+ \frac{\alpha_3^2}{(a_1(1-a_3) - a_2^2)} (\hat{a}_3^* - a_3^*) + O\left(\frac{1}{T}\right).
\]
(A.12)

The canonical matrices of \(a_1\), \(a_2\) and \(a_3\) can be expanded in the same way as \(\hat{\theta}\) in the Theorem I, as shown in the Internet Appendix. Also, \(\hat{a}, \hat{b}, \hat{c}, \hat{a}^*, \hat{b}^*\) and \(\hat{c}^*\) can be similarly expanded and the canonical matrices are:

\[
C_a = -\frac{\alpha_3^2}{(a_1(1-a_3) - a_2^2)} C_{a_1} + \frac{2\alpha_2\alpha_3}{(a_1(1-a_3) - a_2^2)} C_{a_2} + \frac{\alpha_3^2}{(a_1(1-a_3) - a_2^2)} C_{a_3},
\]
(A.13)

\[
C_{a^*} = -\frac{(a_2^*)^2}{(a_1(1-a_3) - a_2^*)^2} C_{a_1^*} + \frac{2\alpha_2\alpha_3}{(a_1(1-a_3) - a_2^*)^2} C_{a_2^*} + \frac{(a_1^*)^2}{(a_1(1-a_3) - a_2^*)^2} C_{a_3^*},
\]
(A.14)

\[
D_a = -\frac{\alpha_3^2}{(a_1(1-a_3) - a_2^2)} D_{a_1} + \frac{2\alpha_2\alpha_3}{(a_1(1-a_3) - a_2^2)} D_{a_2} + \frac{\alpha_3^2}{(a_1(1-a_3) - a_2^2)} D_{a_3},
\]
(A.15)

\[
D_{a^*} = -\frac{(a_2^*)^2}{(a_1(1-a_3) - a_2^*)^2} D_{a_1^*} + \frac{2\alpha_2\alpha_3}{(a_1(1-a_3) - a_2^*)^2} D_{a_2^*} + \frac{(a_1^*)^2}{(a_1(1-a_3) - a_2^*)^2} D_{a_3^*},
\]
(A.16)

\[
C_b = -\frac{\alpha_3(1-a_3)}{(a_1(1-a_3) - a_2^2)} C_{b_1} + \frac{\alpha_1(1-a_3)+a_2^2}{(a_1(1-a_3) - a_2^2)} C_{b_2} + \frac{\alpha_3\alpha_2}{(a_1(1-a_3) - a_2^2)} C_{b_3},
\]
(A.17)

\[
C_{b^*} = -\frac{\alpha_3(1-a_3)}{(a_1(1-a_3) - a_2^*)^2} C_{b_1^*} + \frac{\alpha_1(1-a_3)+a_2^*}{(a_1(1-a_3) - a_2^*)^2} C_{b_2^*} + \frac{\alpha_3\alpha_2}{(a_1(1-a_3) - a_2^*)^2} C_{b_3^*},
\]
(A.18)

\[
D_b = -\frac{\alpha_3(1-a_3)}{(a_1(1-a_3) - a_2^2)} D_{b_1} + \frac{\alpha_1(1-a_3)+a_2^2}{(a_1(1-a_3) - a_2^2)} D_{b_2} + \frac{\alpha_3\alpha_2}{(a_1(1-a_3) - a_2^2)} D_{b_3},
\]
(A.19)
\[
D_{b^*} = -\frac{\alpha_2^*}{\alpha_1^*(1+\alpha_2^*)} D_{\alpha_1^*} + \frac{\alpha_2^*(1+\alpha_2^*) - (\alpha_2^*)^2}{(\alpha_1^*(1+\alpha_2^*) - (\alpha_2^*)^2)} D_{\alpha_2^*} + \frac{\alpha_2^*}{(\alpha_1^*(1+\alpha_2^*) - (\alpha_2^*)^2)^2} D_{\alpha_3^*}, \quad (A.20)
\]

\[
C_c = -\frac{(1+\alpha_3)^2}{(\alpha_1^*(1+\alpha_3) - \alpha_2^*)^2} C_{\alpha_1^*} + \frac{2\alpha_2^*(1+\alpha_3)}{(\alpha_1^*(1+\alpha_3) - \alpha_2^*)^2} C_{\alpha_2^*} + \frac{\alpha_2^*}{(\alpha_1^*(1+\alpha_3) - \alpha_2^*)^2} C_{\alpha_3^*}, \quad (A.21)
\]

\[
C_{c^*} = -\frac{(1+\alpha_3)^2}{(\alpha_1^*(1+\alpha_3) - \alpha_2^*)^2} C_{\alpha_1^*} + \frac{2\alpha_2^*(1+\alpha_3)}{(\alpha_1^*(1+\alpha_3) - \alpha_2^*)^2} C_{\alpha_2^*} + \frac{(\alpha_2^*)^2}{(\alpha_1^*(1+\alpha_3) - \alpha_2^*)^2} C_{\alpha_3^*}, \quad (A.22)
\]

\[
D_c = -\frac{(1+\alpha_3)^2}{(\alpha_1^*(1+\alpha_3) - \alpha_2^*)^2} D_{\alpha_1^*} + \frac{2\alpha_2^*(1+\alpha_3)}{(\alpha_1^*(1+\alpha_3) - \alpha_2^*)^2} D_{\alpha_2^*} + \frac{\alpha_2^*}{(\alpha_1^*(1+\alpha_3) - \alpha_2^*)^2} D_{\alpha_3^*}, \quad (A.23)
\]

\[
D_{c^*} = -\frac{(1+\alpha_3)^2}{(\alpha_1^*(1+\alpha_3) - \alpha_2^*)^2} D_{\alpha_1^*} + \frac{2\alpha_2^*(1+\alpha_3)}{(\alpha_1^*(1+\alpha_3) - \alpha_2^*)^2} D_{\alpha_2^*} + \frac{(\alpha_2^*)^2}{(\alpha_1^*(1+\alpha_3) - \alpha_2^*)^2} D_{\alpha_3^*}. \quad (A.24)
\]

**Case 2:** Test when one Squared Sharpe ratio is for a maximum correlation portfolio with respect to \( Z \).

The squared Sharpe ratio of the maximum correlation portfolio, \( p \), can be written as \( S_p^2 = \frac{\mu_p^2 - 2\mu_p \Phi + \Phi^2}{\sigma_p^2} \). Define

\[
a \equiv \frac{\mu_p^2}{\sigma_p^2}, \quad b \equiv \frac{\mu_p}{\sigma_p} \quad \text{and} \quad c \equiv \frac{1}{\sigma_p^2}. \]

Following the derivation for case 1,

\[
\hat{a} - a = \frac{2\mu_p}{\sigma_p^2} \left( \bar{\mu}_p - \mu_p \right) - \left( \frac{\mu_p^2}{\sigma_p^2} \right) \left( \bar{\sigma}_p^2 - \sigma_p^2 \right) + O\left( \frac{1}{p} \right), \quad (A.25)
\]

\[
\hat{b} - b = \frac{1}{\sigma_p} \left( \bar{\mu}_p - \mu_p \right) - \frac{\mu_p}{\sigma_p^2} \left( \bar{\sigma}_p^2 - \sigma_p^2 \right) + O\left( \frac{1}{p} \right), \quad (A.26)
\]

\[
\hat{c} - c = -\frac{1}{\sigma_p^2} \left( \bar{\sigma}_p^2 - \sigma_p^2 \right) + O\left( \frac{1}{p} \right). \quad (A.27)
\]

From equation (A.113) and (A.114),

\[
C_a = \frac{2\mu_p}{\sigma_p^2} C_{\mu p} - \left( \frac{\mu_p^2}{\sigma_p^2} \right) C_{\sigma p^2}, \quad (A.28)
\]

\[
C_b = \frac{1}{\sigma_p^2} C_{\mu p} - \frac{\mu_p}{\sigma_p^2} C_{\sigma p^2}, \quad (A.29)
\]

\[
C_c = -\frac{1}{\sigma_p} C_{\sigma p^2}, \quad (A.30)
\]

\[
D_a = \frac{2\mu_p}{\sigma_p^2} D_{\mu p} - \left( \frac{\mu_p^2}{\sigma_p^2} \right) D_{\sigma p^2}, \quad (A.31)
\]

\[
D_b = \frac{1}{\sigma_p^2} D_{\mu p} - \frac{\mu_p}{\sigma_p^2} D_{\sigma p^2}, \quad (A.32)
\]

\[
D_c = -\frac{1}{\sigma_p^2} D_{\sigma p^2}. \quad (A.33)
\]
The other squared Sharpe ratios can be written as $\alpha^* - 2b^* \phi + c^* \phi^2$, and $\hat{\alpha}^*, \hat{\beta}^*$ and $\hat{\epsilon}^*$ can be expanded in the same way as $\hat{\theta}$ in the Theorem I, with canonical matrices $C_\alpha^*, D_\alpha^*, C_\beta^*, D_\beta^*, C_c^*$ and $D_c^*$. Then the asymptotic variance of BS test with a common optimal zero-beta rate also follows the same procedure in case 1.

Case 3: Tests where one of the Squared Sharpe ratios is for a fixed weight portfolio.

Similar to case 2, the squared Sharpe ratio of the maximum correlation portfolio can be written as $S_\rho^2 = \mu_\rho^2 - 2\mu_\rho \phi + \phi^2$. Define $a \equiv \mu_\rho^2 / \sigma_\rho^2$, $b \equiv \mu_\rho / \sigma_\rho$ and $c \equiv 1 / \sigma_\rho^2$. The estimates for the canonical matrices are

\[ C_\mu_p = \frac{A^{-1}}{T} \left( \sum_{t=1}^{T} Z_{t-1} \right) w', \quad D_\mu_p = 0, \]  \[ (A.34) \]

\[ C_{\sigma_p^2} = \frac{2A^{-1}}{T} \sum_{t=1}^{T} Z_{t-1} \left[ - (\mu_t w - \mu_p) \right] w', \quad D_{\sigma_p^2} = -ww'. \]  \[ (A.35) \]

Proof of the Theorem I.

We first express $\hat{\theta}$ in terms of $\epsilon_t$ as follows:

\[ \hat{\theta} = \theta + \sum_{t=1}^{T} C_t \left( \frac{1}{T} \sum_{t=1}^{T} \epsilon_t Z_{t-1}' \right) A^{-1} Z_{t-1} + tr \left[ D \frac{1}{T} \sum_{t=1}^{T} (\epsilon_t \epsilon_t' - V) \right] + O_p \left( 1/T \right) \]  \[ (A.36) \]

where we have used the fact that $\hat{\mu}_t - \mu_t = (\hat{\delta} - \delta)' Z_{t-1} = \left( \frac{1}{T} \sum_{t=1}^{T} \epsilon_t Z_{t-1}' \right) A^{-1} Z_{t-1} = O_p \left( 1/\sqrt{T} \right)$ which follows from (13), and the fact that $\hat{V} - V = \frac{1}{T} \sum_{t=1}^{T} (\epsilon_t \epsilon_t' - V) + O_p \left( 1/T \right)$. Then

\[ \frac{1}{T} \sum_{t=1}^{T} \hat{\epsilon} \hat{\epsilon}' = \frac{1}{T} \sum_{t=1}^{T} \left[ \epsilon_t - (\hat{\delta} - \delta)' Z_{t-1} \right] \left[ \epsilon_t - (\hat{\delta} - \delta)' Z_{t-1} \right]' = \frac{1}{T} \sum_{t=1}^{T} \epsilon_t \epsilon_t' - (\hat{\delta} - \delta)' \left( \frac{1}{T} \sum_{t=1}^{T} Z_{t-1} \epsilon_t' \right) - (\hat{\delta} - \delta)' A(\hat{\delta} - \delta) + O_p \left( 1/T \right) \]

Using the $L \times N$ matrix (without subscript) $C = A^{-1} \sum_{t=1}^{T} Z_{t-1} C_t'$, we may rewrite (A.1) using matrix commutativity within the trace operator as follows:
\[ \hat{\theta} = \theta + tr \left[ \sum_{i=1}^{T} A^{-1} Z_{t-1} C_{t} \left( \frac{1}{T} \sum_{i=1}^{T} \varepsilon_{t} Z'_{t-1} \right) \right] + tr \left[ D \frac{1}{T} \sum_{i=1}^{T} (\varepsilon_{t} \varepsilon'_{t} - V) \right] + O_p \left( \frac{1}{T} \right) \]

\[ = \theta + tr \left[ C \left( \frac{1}{T} \sum_{i=1}^{T} \varepsilon_{t} Z'_{t-1} \right) \right] + tr \left[ \frac{1}{T} \sum_{i=1}^{T} (D \varepsilon_{t} \varepsilon'_{t} - DV) \right] + O_p \left( \frac{1}{T} \right) \]

\[ = \theta + \frac{1}{T} \sum_{i=1}^{T} (Z'_{t-1} C \varepsilon_{t} + \varepsilon'_{t} D \varepsilon_{t}) - tr(DV) + O_p \left( \frac{1}{T} \right) \]

(A.37)

Where \( C, D \) and \( V \) are the true values of the coefficients.

Now we let \( \psi_{t} = Z_{t-1}' C \varepsilon_{t} + \varepsilon_{t}' D \varepsilon_{t} \). The unconditional asymptotic variance of \( \hat{\theta} \) is

\[ \sqrt{T} \text{Var}\left\{ \left( \frac{1}{T} \sum_{t=1}^{T} (Z_{t-1}' C \varepsilon_{t} + \varepsilon_{t}' D \varepsilon_{t}) \right) \right\} \] and this may be estimated simply as:

\[ \sqrt{T} \left\{ \left( \frac{1}{T} \sum_{t=1}^{T} \psi_{t} \right)^2 - \left[ \left( \frac{1}{T} \sum_{t=1}^{T} \psi_{t} \right) \right]^2 \right\} \] (A.38)

This follows because \( \psi_{t} \) has a moving average structure of order zero, as may be seen by examining for any \( j \geq 1 \) the autocovariance, using the independence of \( \varepsilon_{t} \):

\[
\text{Cov}(\psi_{t}, \psi_{t-j}) = \text{Cov}(Z_{t-1}' C \varepsilon_{t} + \varepsilon_{t}' D \varepsilon_{t} ; Z_{t-j-2}' C \varepsilon_{t-j-1} + \varepsilon_{t-j-1}' D \varepsilon_{t-j-1})
\]

\[ = \text{Cov}(Z_{t-1}' C \varepsilon_{t} ; Z_{t-j-2}' C \varepsilon_{t-j-1}) \]

\[ = E\{ Z_{t-1}' C E(\varepsilon_{t} | \varepsilon_{t-j-1}) \varepsilon_{t-j-1}' C' Z_{t-j-2} \} = 0. \]

Substituting consistent estimates \( \hat{C}, \hat{D} \) along with \( \hat{\varepsilon}_{t} = R_{t} - \hat{\delta}' Z_{t-1} \) to obtain a consistent estimator of \( V \), we have a consistent estimator for the asymptotic variance. \( \square \)

Note that it is possible to find the asymptotic variance of a combination of two estimators by using the \( \hat{C} \) and \( \hat{D} \) matrices associated with each estimator. For example, the asymptotic variance of the sum of two estimators, \( \hat{\theta} + \hat{\theta}' \), is characterized by matrices \( \hat{C} + \hat{C}' \) and \( \hat{D} + \hat{D}' \) due to linearity of (A.1),
allowing us to combine the estimators' \( \hat{C} \) and \( \hat{D} \) matrices. The asymptotic variance of the ratio \( \hat{\theta}/\hat{\theta}^* \) is characterized by matrices \( \frac{\hat{\theta} \hat{C} - \hat{\theta} \hat{C}^*}{\hat{\theta}^2} \) and \( \frac{\hat{\theta} \hat{D} - \hat{\theta} \hat{D}^*}{\hat{\theta}^2} \), as may be seen from the following expansion:

\[
\frac{\hat{\theta}}{\hat{\theta}^*} = \frac{\theta + \sum_{t=1}^{T} C_t'(\hat{\mu}_t - \mu_t) + tr\left[D(\hat{V} - V)\right] + O_p(1/T)}{\theta + \sum_{t=1}^{T} C''_t(\hat{\mu}_t - \mu_t) + tr\left[D'(\hat{V} - V)\right] + O_p(1/T)}
\]

\[
= \frac{\theta}{\theta^*} \left[1 + \frac{1}{\theta} \sum_{t=1}^{T} C_t'(\hat{\mu}_t - \mu_t) + \frac{1}{\theta} tr\left[D(\hat{V} - V)\right] + O_p(1/T)\right] \times
\]

\[
\left[1 - \frac{1}{\theta} \sum_{t=1}^{T} C''_t(\hat{\mu}_t - \mu_t) - \frac{1}{\theta} tr\left[D'(\hat{V} - V)\right] + O_p(1/T)\right]
\]

\[
= \frac{\theta}{\theta^*} \sum_{t=1}^{T} \frac{\theta C_t' - \theta C''_t}{\theta^2} (\hat{\mu}_t - \mu_t) + tr\left[\frac{\theta D - \theta D^*}{\theta^2} (\hat{V} - V)\right] + O_p(1/T)
\]

The asymptotic variance of \( \frac{\hat{\theta} - \hat{\theta}^*}{1 + \hat{\theta}^*} \) is characterized by the matrices \( \left[(1 + \hat{\theta}^*) \hat{C} - (1 + \hat{\theta}) \hat{C}^*\right]/(1 + \hat{\theta})^2 \) and \( \left[(1 + \hat{\theta}^*) \hat{D} - (1 + \hat{\theta}) \hat{D}^*\right]/(1 + \hat{\theta})^2 \), as may be seen from the following expansion:

\[
\frac{\hat{\theta} - \hat{\theta}^*}{1 + \hat{\theta}^*} = \frac{\theta - \theta^*}{1 + \theta^*} \sum_{t=1}^{T} \frac{C_t' - C''_t}{\theta - \theta^*} (\hat{\mu}_t - \mu_t) + tr\left[(D - D^*)(\hat{V} - V)\right] + O_p(1/T)
\]

\[
= \frac{\theta - \theta^*}{1 + \theta^*} \left[1 + \sum_{t=1}^{T} \frac{C_t' - C''_t}{\theta - \theta^*} (\hat{\mu}_t - \mu_t) + tr\left[D - D^*(\hat{V} - V)\right] + O_p(1/T)\right] \times
\]

\[
\left[1 - \sum_{t=1}^{T} \frac{C''_t}{\theta - \theta^*} (\hat{\mu}_t - \mu_t) - tr\left[D^*/(1 + \theta^*) (\hat{V} - V)\right] + O_p(1/T)\right]
\]

\[
= \frac{\theta - \theta^*}{1 + \theta^*} \left[1 + \sum_{t=1}^{T} \frac{C_t' - C''_t}{\theta - \theta^*} (\hat{\mu}_t - \mu_t) + tr\left[D - D^*/(1 + \theta^*) (\hat{V} - V)\right] + O_p(1/T)\right]
\]

\[
= \frac{\theta - \theta^*}{1 + \theta^*} \sum_{t=1}^{T} \frac{(1 + \theta^*) C_t' - (1 + \theta) C''_t}{(1 + \theta)^2} (\hat{\mu}_t - \mu_t) + tr\left[\frac{(1 + \theta^*) D - (1 + \theta) D^*}{(1 + \theta)^2} (\hat{V} - V)\right] + O_p(1/T)
\]
Proof of Proposition II:

From Ferson & Siegel (2001, pages 976-977) the weights of the UE portfolio with mean $\mu_p$, in the absence of a riskless rate, are

$$w'(Z) = \frac{1'\Lambda(Z)}{1'\Lambda(Z)1'} + \frac{\mu_p - \alpha_2}{1 - \alpha_3} \mu'(Z) \left( \Lambda(Z) - \frac{\Lambda(Z)1'\Lambda(Z)}{1'\Lambda(Z)1'} \right)$$  \hfill (A.41)

where our $\alpha_3$ represents $1 - \alpha_3$ from Ferson and Siegel (2001). We will write with implicit $Z_{t-1}$ (the conditioning information) dependence as

$$w' = \frac{1'\Lambda}{1'\Lambda 1'} + \frac{\mu_p - \alpha_2}{1 - \alpha_3} \mu' \left( \Lambda - \frac{\Lambda 1'\Lambda}{1'\Lambda 1'} \right)$$  \hfill (A.42)

where

$$\Lambda = \Lambda(Z) = \left[ V + \mu(Z)\mu'(Z) \right]^{-1} = (V + \mu\mu')^{-1}$$  \hfill (A.43)

The portfolio variance is

$$\sigma_p^2 = \left( \alpha_1 + \frac{\alpha_2^2}{1 - \alpha_3} \right) - \frac{2\alpha_2}{1 - \alpha_3} + \frac{\alpha_3}{1 - \alpha_3}$$  \hfill (A.44)

To simplify notation in what follows, we define the following:

$$a = 1'V^{-1}1, \quad b_i = 1'V^{-1}\mu_i, \quad c_{ki} = 1 + \mu_k'V^{-1}\mu_i, \quad d_{ki} = ac_{ki} - b_k b_i$$  \hfill (A.45)

along with their estimates

$$\hat{a} = 1'\hat{V}^{-1}1, \quad \hat{b}_i = 1'\hat{V}^{-1}\hat{\mu}_i, \quad \hat{c}_{ki} = 1 + \hat{\mu}_k'\hat{V}^{-1}\hat{\mu}_i, \quad \hat{d}_{ki} = \hat{a}\hat{c}_{ki} - \hat{b}_k \hat{b}_i$$  \hfill (A.46)

Consistent estimates of the portfolio constants are then:
\[ \hat{\alpha}_1 = \frac{1}{T} \sum_{t=1}^{T} \frac{\hat{c}_t}{\hat{d}_t}, \quad \hat{\alpha}_2 = \frac{1}{T} \sum_{t=1}^{T} \frac{\hat{b}_t}{\hat{d}_t}, \quad \text{and} \quad \hat{\alpha}_3 = \frac{\hat{a}}{T} \sum_{t=1}^{T} \frac{1}{\hat{d}_t} \]  

(A.47)

Here we restate the proposition for convenience.

**Proposition 11**: The estimation bias of the estimated maximized squared Sharpe Ratio at zero-beta rate \( \varphi \):

\[
\hat{S}_\varphi^2 = \frac{\hat{\alpha}_1 - 2\varphi\hat{\alpha}_2 + \varphi^2\hat{\alpha}_3}{\hat{\alpha}_1\hat{\alpha}_3 - \hat{\alpha}_2^2} - 1
\]

(A.48)

with respect to the true (but unknown) maximized squared Sharpe Ratio

\[
S^2_\varphi = \frac{\alpha_1 - 2\varphi\alpha_2 + \varphi^2\alpha_3}{\alpha_1\alpha_3 - \alpha_2^2} - 1
\]

(A.49)

may be expressed asymptotically by taking the expectation of its second-order Taylor Series expansion:

\[
Bias = E\left(\hat{S}_\varphi^2 - S^2_\varphi\right) \equiv \left(\sum_{i=1}^{3} E\left(\hat{\alpha}_i - \alpha_i\right) \frac{\partial S^2_\varphi}{\partial \alpha_i}\right) + \left(\sum_{i,j=1}^{3} \frac{E\left[(\hat{\alpha}_i - \alpha_i)(\hat{\alpha}_j - \alpha_j)\right]}{2} \frac{\partial^2 S^2_\varphi}{\partial \alpha_i \partial \alpha_j}\right)
\]

(A.50)

for which the partial derivatives are derived in Proposition 2, with the expectation terms following in Propositions 3a, 3b, and 3c. Note that the estimated bias may be obtained using estimates in the formulas for derivatives and expectations that follow.

The partial derivatives for Proposition 1 are as follows:

\[
\frac{\partial S^2_\varphi}{\partial \alpha_1} = -\left(\frac{\alpha_2 - \varphi\alpha_3}{\alpha_1\alpha_3 - \alpha_2^2}\right)^2, \quad \frac{\partial S^2_\varphi}{\partial \alpha_2} = 2\left(\frac{\alpha_1 - \varphi\alpha_2}{\alpha_1\alpha_3 - \alpha_2^2}\right)\left(\frac{\alpha_2 - \varphi\alpha_1}{\alpha_1\alpha_3 - \alpha_2^2}\right), \quad \frac{\partial S^2_\varphi}{\partial \alpha_3} = -\left(\frac{\alpha_1 - \varphi\alpha_2}{\alpha_1\alpha_3 - \alpha_2^2}\right)^2,
\]
\[
\frac{\partial^2 S^2_\varphi}{\partial \alpha_i} = 2 \alpha_i \left(\alpha_2 - \varphi \alpha_3\right)^2, \quad \frac{\partial^2 S^2_\varphi}{\partial \alpha_2} = 2 \alpha_1 - 2 \varphi \alpha_2 + \varphi^2 \alpha_3, \quad \frac{\partial^2 S^2_\varphi}{\partial \alpha_3} = 2 \alpha_2 - 2 \varphi \alpha_3 + \varphi^2 \alpha_i, \quad \frac{\partial^2 S^2_\varphi}{\partial \alpha_i \partial \alpha_2} = \frac{\alpha_1}{\varphi} \left(\alpha_1, \alpha_3 - \alpha^2, \alpha_2\right) + 8 \alpha_2 \left(\alpha_1 - \varphi \alpha_2\right) \left(\alpha_3 - \alpha^2\right) + \left(\alpha_1 - \varphi \alpha_2\right), \quad \frac{\partial^2 S^2_\varphi}{\partial \alpha_i \partial \alpha_3} = -\frac{\alpha_2 \alpha_3}{\varphi} \left(\alpha_1, \alpha_3 - \alpha^2, \alpha_3\right) + \left(\alpha_1 - \varphi \alpha_2\right) \left(\alpha_3 - \alpha^2\right), \quad (A.51)
\]

Proof: The expression for the approximate asymptotic bias represents the second-order Taylor Series. As in Siegel and Woodgate (2007) “we use the method of statistical differentials to find Taylor-series approximations to expectations of random variables, obtaining results that are asymptotically correct when the number of time periods is large and that remain statistically consistent when estimated values are substituted for unknown parameters.” Note that the \( \hat{\alpha}_i \) are not necessarily unbiased, so their expectations are necessary in this expression, concluding the proof. Partial differentiation shows:

\[\begin{align*}
\frac{\partial S^2_\varphi}{\partial \alpha_1} &= \frac{\alpha_1 \alpha_3 - \alpha_2}{\left(\alpha_1, \alpha_3 - \alpha^2, \alpha_2\right)^2} \left(\alpha_1 - 2 \varphi \alpha_2 + \varphi^2 \alpha_3\right) - \alpha_2 \varphi \alpha_3 + \varphi^2 \alpha^3 \left(\alpha_1 - \varphi \alpha_2\right) \\
&= \frac{\alpha_1 \alpha_3 - \alpha_2 - \alpha_1 \alpha_3 + 2 \varphi \alpha_2 \alpha_3 - \varphi^2 \alpha_3^2}{\left(\alpha_1, \alpha_3 - \alpha^2, \alpha_3\right)^2} - \frac{\alpha_2 \varphi \alpha_3 - \varphi^2 \alpha_3^2}{\left(\alpha_1, \alpha_3 - \alpha^2\right)^2} \\
&= -\frac{\alpha_2 - \varphi \alpha_3}{\left(\alpha_1, \alpha_3 - \alpha^2\right)^2}
\end{align*}\]
\[
\frac{\partial S^2_\varphi}{\partial \alpha_2} = \frac{(-2\varphi)(\alpha_1, \alpha_3 - \alpha_2^2) - (\alpha_1 - 2\varphi \alpha_2 + \varphi^2 \alpha_3)(-2\alpha_2)}{(\alpha_1, \alpha_3 - \alpha_2^2)^2} \\
= 2 \frac{-\varphi \alpha_1, \alpha_3 - \alpha_2^2 + \alpha_1, \alpha_2 - 2\varphi \alpha_2^2 + \varphi^2 \alpha_2 \alpha_3}{(\alpha_1, \alpha_3 - \alpha_2^2)^2} - \frac{2 \alpha, \alpha_2 - \varphi \alpha_1, \alpha_3 - \alpha_2^2 + \varphi^2 \alpha_2 \alpha_3}{(\alpha_1, \alpha_3 - \alpha_2^2)^2} \\
= 2 \frac{\alpha_1, \alpha_2 - \varphi \alpha_1, \alpha_3 - \varphi \alpha_2 (\alpha_2 - \varphi \alpha_3)}{(\alpha_1, \alpha_3 - \alpha_2^2)^2} - \frac{2 \alpha_1, \alpha_2 - \varphi \alpha_1, \alpha_3}{(\alpha_1, \alpha_3 - \alpha_2^2)^2} \\
= 2 \left( \frac{\alpha_1 - \varphi \alpha_2}{\alpha_1, \alpha_3 - \alpha_2^2} \right)^2 \\
\]

\[
\frac{\partial \varphi^2}{\partial \alpha_3} = \frac{\varphi}{(\alpha_1, \alpha_3 - \alpha_2^2)^2} - \frac{(\alpha_1 - 2\varphi \alpha_2 + \varphi^2 \alpha_3)}{\alpha_1, \alpha_3 - \alpha_2^2} \\
= \frac{\varphi^2 \alpha_1 - \alpha_2^2 - \alpha_2^2 + \varphi \alpha \alpha_2 - \varphi^2 \alpha_3}{(\alpha_1, \alpha_3 - \alpha_2^2)^2} \\
= -\frac{\varphi^2 \alpha_2^2 - \alpha_1^2 + 2\varphi \alpha \alpha_2}{(\alpha_1, \alpha_3 - \alpha_2^2)^2} = -\left( \frac{\alpha_1 - \varphi \alpha_2}{\alpha_1, \alpha_3 - \alpha_2^2} \right)^2 \\
\]

\[
\frac{\partial^2 S^2_\varphi}{\partial \alpha_1^2} = -\frac{\partial}{\partial \alpha_1} \left( \frac{\alpha_2 - \varphi \alpha_3}{\alpha_1, \alpha_3 - \alpha_2^2} \right)^2 \\
= -\frac{\partial}{\partial \alpha_1} \left( \frac{\alpha_2 - \varphi \alpha_3}{\alpha_1, \alpha_3 - \alpha_2^2} \right)^2 \\
= -\frac{2(\alpha_2 - \varphi \alpha_3)^2 (\alpha_1, \alpha_3 - \alpha_2^2)}{(\alpha_1, \alpha_3 - \alpha_2^2)^4} - \frac{2\alpha_3 (\alpha_2 - \varphi \alpha_3)^2}{(\alpha_1, \alpha_3 - \alpha_2^2)^3} \\
\]
\[
\frac{\partial^2 S^2}{\partial \alpha_2^2} = 2 \frac{\partial}{\partial \alpha_2} \left( \frac{\alpha,\alpha_2 - \varphi \alpha,\alpha_3 - \varphi \alpha_2^2 + \varphi^2 \alpha_2,\alpha_3}{(\alpha,\alpha_3 - \alpha_2^2)^2} \right) \\
= 2 \left( \alpha_1 - 2 \varphi \alpha_2 + \varphi^2 \alpha_3 \right) \frac{(\alpha,\alpha_3 - \alpha_2^2)^2 - (\alpha,\alpha_2 - \varphi \alpha,\alpha_3 - \varphi \alpha_2^2 + \varphi^2 \alpha_2,\alpha_3) 2(\alpha,\alpha_3 - \alpha_2^2)(-2\alpha_2)}{(\alpha,\alpha_3 - \alpha_2^2)^4} \\
= 2 \left( \alpha_1 - 2 \varphi \alpha_2 + \varphi^2 \alpha_3 \right) \frac{(\alpha,\alpha_3 - \alpha_2^2) + 4(\alpha,\alpha_2 - \varphi \alpha,\alpha_3 - \varphi \alpha_2^2 + \varphi^2 \alpha_2,\alpha_3) \alpha_2}{(\alpha,\alpha_3 - \alpha_2^2)^3} \\
= 2 \left( \alpha_1 - 2 \varphi \alpha_2 + \varphi^2 \alpha_3 \right) \frac{(\alpha,\alpha_3 - \alpha_2^2)^2 + 8(\alpha,\alpha_2 - \varphi \alpha,\alpha_3 - \varphi \alpha_2^2 + \varphi^2 \alpha_2,\alpha_3) \alpha_2}{(\alpha,\alpha_3 - \alpha_2^2)^3} \\
= 2 \alpha_1 - 2 \varphi \alpha_2 + \varphi^2 \alpha_3 + 8 \frac{\alpha_1,\alpha_2 - \varphi \alpha,\alpha_3 - \varphi \alpha_2^2 + \varphi^2 \alpha_2,\alpha_3}{(\alpha,\alpha_3 - \alpha_2^2)^2} \\
= 2 \alpha_1 - 2 \varphi \alpha_2 + \varphi^2 \alpha_3 + 8 \frac{\alpha_1,\alpha_2 - \varphi \alpha,\alpha_3 - \varphi \alpha_2^2 + \varphi^2 \alpha_2,\alpha_3}{(\alpha,\alpha_3 - \alpha_2^2)^3} \\
= 2 \alpha_1 - 2 \varphi \alpha_2 + \varphi^2 \alpha_3 + 8 \frac{\alpha_1,\alpha_2 - \varphi \alpha,\alpha_3 - \varphi \alpha_2^2 + \varphi^2 \alpha_2,\alpha_3}{(\alpha,\alpha_3 - \alpha_2^2)^3} \\
\]

\[
\frac{\partial^2 S^2}{\partial \alpha_3^2} = \frac{\partial}{\partial \alpha_3} \left[ -\frac{(\alpha_1 - \varphi \alpha_2)^2}{(\alpha,\alpha_3 - \alpha_2^2)^2} \right] = -\frac{(\alpha_1 - \varphi \alpha_2)^2}{(\alpha,\alpha_3 - \alpha_2^2)^4} \frac{2(\alpha,\alpha_3 - \alpha_2^2)(\alpha_1) - 2 \alpha_1 (\alpha_1 - \varphi \alpha_2)^2}{(\alpha,\alpha_3 - \alpha_2^2)^3} 
\]
\[ \frac{\partial^2 S_\phi^2}{\partial \alpha_1 \partial \alpha_2} = -\frac{\partial}{\partial \alpha_2} \left( \frac{\alpha_2 - \varphi \alpha_3}{\alpha_1 \alpha_3 - \alpha_2^2} \right)^2 = -\frac{\partial}{\partial \alpha_2} \left( \frac{\alpha_2 - \varphi \alpha_3}{\alpha_1 \alpha_3 - \alpha_2^2} \right)^2 \\
= -2 \left( \alpha_2 - \varphi \alpha_3 \right) \left( \alpha_1 \alpha_3 - \alpha_2^2 \right)^2 \left( \alpha_2 - \varphi \alpha_3 \right)^2 2 \left( \alpha_1 \alpha_3 - \alpha_2 \right) (-2 \alpha_2) \\
= -2 \left( \alpha_2 - \varphi \alpha_3 \right) \left( \alpha_1 \alpha_3 - \alpha_2^2 \right)^3 + 2 \left( \alpha_2 - \varphi \alpha_3 \right)^2 \alpha_2 \\
= -2 \left( \alpha_2 - \varphi \alpha_3 \right) \left( \alpha_1 \alpha_3 - \alpha_2^2 + 2 \left( \alpha_2 - \varphi \alpha_3 \right) \alpha_2 \right) \\
= -2 \left( \alpha_2 - \varphi \alpha_3 \right) \left( \alpha_1 \alpha_3 - \alpha_2^2 + 2 \left( \alpha_2 - \varphi \alpha_3 \right) \alpha_2 \right) \\
= -2 \left( \alpha_2 - \varphi \alpha_3 \right) \left( \alpha_1 \alpha_3 - \alpha_2^2 + 2 \left( \alpha_2 - \varphi \alpha_3 \right) \alpha_2 \right) \\
= -2 \left( \alpha_2 - \varphi \alpha_3 \right) \left( \alpha_1 \alpha_3 - \alpha_2^2 + 2 \left( \alpha_2 - \varphi \alpha_3 \right) \alpha_2 \right) \\
= -2 \left( \alpha_2 - \varphi \alpha_3 \right) \left( \alpha_1 \alpha_3 - \alpha_2^2 + 2 \left( \alpha_2 - \varphi \alpha_3 \right) \alpha_2 \right) \\
= -2 \left( \alpha_2 - \varphi \alpha_3 \right) \left( \alpha_1 \alpha_3 - \alpha_2^2 + 2 \left( \alpha_2 - \varphi \alpha_3 \right) \alpha_2 \right) \\
= -2 \left( \alpha_2 - \varphi \alpha_3 \right) \left( \alpha_1 \alpha_3 - \alpha_2^2 + 2 \left( \alpha_2 - \varphi \alpha_3 \right) \alpha_2 \right) \\
= -2 \left( \alpha_2 - \varphi \alpha_3 \right) \left( \alpha_1 \alpha_3 - \alpha_2^2 + 2 \left( \alpha_2 - \varphi \alpha_3 \right) \alpha_2 \right) \\
= -2 \left( \alpha_2 - \varphi \alpha_3 \right) \left( \alpha_1 \alpha_3 - \alpha_2^2 + 2 \left( \alpha_2 - \varphi \alpha_3 \right) \alpha_2 \right) \\
= -2 \left( \alpha_2 - \varphi \alpha_3 \right) \left( \alpha_1 \alpha_3 - \alpha_2^2 + 2 \left( \alpha_2 - \varphi \alpha_3 \right) \alpha_2 \right) \\
= -2 \left( \alpha_2 - \varphi \alpha_3 \right) \left( \alpha_1 \alpha_3 - \alpha_2^2 + 2 \left( \alpha_2 - \varphi \alpha_3 \right) \alpha_2 \right) \\
= -2 \left( \alpha_2 - \varphi \alpha_3 \right) \left( \alpha_1 \alpha_3 - \alpha_2^2 + 2 \left( \alpha_2 - \varphi \alpha_3 \right) \alpha_2 \right) \\
= -2 \left( \alpha_2 - \varphi \alpha_3 \right) \left( \alpha_1 \alpha_3 - \alpha_2^2 + 2 \left( \alpha_2 - \varphi \alpha_3 \right) \alpha_2 \right) \\
= -2 \left( \alpha_2 - \varphi \alpha_3 \right) \left( \alpha_1 \alpha_3 - \alpha_2^2 + 2 \left( \alpha_2 - \varphi \alpha_3 \right) \alpha_2 \right) \\
= -2 \left( \alpha_2 - \varphi \alpha_3 \right) \left( \alpha_1 \alpha_3 - \alpha_2^2 + 2 \left( \alpha_2 - \varphi \alpha_3 \right) \alpha_2 \right) \\
= -2 \left( \alpha_2 - \varphi \alpha_3 \right) \left( \alpha_1 \alpha_3 - \alpha_2^2 + 2 \left( \alpha_2 - \varphi \alpha_3 \right) \alpha_2 \right) \\
= -2 \left( \alpha_2 - \varphi \alpha_3 \right) \left( \alpha_1 \alpha_3 - \alpha_2^2 + 2 \left( \alpha_2 - \varphi \alpha_3 \right) \alpha_2 \right) \\
= -2 \left( \alpha_2 - \varphi \alpha_3 \right) \left( \alpha_1 \alpha_3 - \alpha_2^2 + 2 \left( \alpha_2 - \varphi \alpha_3 \right) \alpha_2 \right) \\
= -2 \left( \alpha_2 - \varphi \alpha_3 \right) \left( \alpha_1 \alpha_3 - \alpha_2^2 + 2 \left( \alpha_2 - \varphi \alpha_3 \right) \alpha_2 \right)
\[
\frac{\delta^2 \delta^2}{\delta \alpha_2 \delta \alpha_3} = -\frac{\partial}{\partial \alpha_2} \left( \frac{\alpha_4 - \varphi \alpha_2^2}{\alpha_1 \alpha_3 - \alpha_2^2} \right)^2 = -2 \left( \frac{\alpha_4 - \varphi \alpha_2^2}{\alpha_1 \alpha_3 - \alpha_2^2} \right) \left( -\varphi \right) \left( \frac{\alpha_4 - \alpha_2}{\alpha_1 \alpha_3 - \alpha_2^2} \right) \left( -\alpha_1 - \varphi \alpha_2 \right) \left( -2 \alpha_2 \right) \left( \alpha_1 \alpha_3 - \alpha_2^2 \right)^2
\]

\[
= 2 \left( \alpha_4 - \varphi \alpha_2^2 \right) \left( \frac{\alpha_4 - \alpha_2}{\alpha_1 \alpha_3 - \alpha_2^2} \right) - 2 \alpha_2 \left( \alpha_1 - \varphi \alpha_2 \right) \left( \frac{\alpha_4 - \alpha_2}{\alpha_1 \alpha_3 - \alpha_2^2} \right)^3
\]

\[
= 2 \left( \alpha_4 - \varphi \alpha_2^2 \right) \left( \frac{\alpha_4 - \alpha_2}{\alpha_1 \alpha_3 - \alpha_2^2} \right) - 4 \left( \alpha_4 - \varphi \alpha_2^2 \right) \left( \frac{\alpha_4 - \alpha_2}{\alpha_1 \alpha_3 - \alpha_2^2} \right)^3
\]

\[
= 2 \left( \frac{\alpha_4 - \varphi \alpha_2^2}{\alpha_1 \alpha_3 - \alpha_2^2} \right) \left( \varphi - 2 \frac{\alpha_4 - \alpha_2}{\alpha_1 \alpha_3 - \alpha_2^2} \right)
\]

(A.52)

Expectations of the form \( E( \hat{\alpha}_i - \alpha_i ) \) are estimated as follows:

\[
E( \hat{\alpha}_1 - \alpha_1 ) \approx \frac{1}{T(T-L)} \sum_{t=1}^T \frac{2a^2 c_n + a \left[ (n-4) c_n - 2 \right] d_n}{d_n^3} \left[ (n-3) + (n-2) c_n \right] d_n^2
\]

\[
+ \frac{1}{T^2} \sum_{t=1}^T \frac{-4a^2 c_n + a \left[ 4 - (n-5) c_n \right] d_n}{d_n^3} \left[ (n-4) d_n^2 \right] Z_{t-1} A^{-1} Z_{t-1}
\]

\[
E( \hat{\alpha}_2 - \alpha_2 ) \approx \frac{1}{T} \sum_{t=1}^T \left( b_n \frac{2a^2 + (n-4) a d_n}{d_n^3 (T-L)} - ab_n \frac{4a + (n-5) d_n}{d_n^3 T} Z_{t-1} A^{-1} Z_{t-1} \right)
\]

and

\[
E( \hat{\alpha}_3 - \alpha_3 ) \approx \frac{a}{T} \sum_{t=1}^T \left( \frac{2a^2 + (n-4) a d_n}{(T-L) d_n^3} - a \frac{4a + (n-5) d_n}{Td_n^3} Z_{t-1} A^{-1} Z_{t-1} \right)
\]

(A.53)

To see this, begin by noting that each of these expectations has the form, expanding to second order:
\[
E \left[ \frac{1}{T} \sum_{t=1}^{T} \left( \frac{\hat{x}_t - x_t}{d_{\alpha}} \right) \right] = \frac{1}{T} \sum_{t=1}^{T} \frac{x_t}{d_{\alpha}} E \left[ \frac{1 + (\hat{x}_t - x_t) / x_t}{1 + (\hat{d}_{\alpha} - d_{\alpha}) / d_{\alpha}} - 1 \right]
\]
\[
\approx \frac{1}{T} \sum_{t=1}^{T} \frac{x_t}{d_{\alpha}} E \left[ \left( 1 + \frac{\hat{x}_t - x_t}{x_t} \right) \left( 1 - \frac{\hat{d}_{\alpha} - d_{\alpha}}{d_{\alpha}} + \left( \frac{\hat{d}_{\alpha} - d_{\alpha}}{d_{\alpha}^2} \right)^2 \right) - 1 \right]
\]
\[
\approx \frac{1}{T} \sum_{t=1}^{T} \frac{x_t}{d_{\alpha}} \left\{ E \left( \frac{\hat{x}_t - x_t}{x_t} \right) - E \left[ \frac{\hat{x}_t - x_t}{x_t} \right] \left( \frac{\hat{d}_{\alpha} - d_{\alpha}}{d_{\alpha}} \right) \right\} - E \left( \frac{\hat{d}_{\alpha} - d_{\alpha}}{d_{\alpha}} \right) + E \left( \frac{\hat{d}_{\alpha} - d_{\alpha}}{d_{\alpha}^2} \right)^2
\]
\[
= \frac{1}{T} \sum_{t=1}^{T} \frac{E(\hat{x}_t - x_t)}{d_{\alpha}} - \frac{1}{T} \sum_{t=1}^{T} E \left( \frac{(\hat{x}_t - x_t)(\hat{d}_{\alpha} - d_{\alpha})}{d_{\alpha}^2} \right) + \frac{1}{T} \sum_{t=1}^{T} x_t \left( \frac{E(\hat{d}_{\alpha} - d_{\alpha})^2}{d_{\alpha}^2} - E \left( \frac{\hat{d}_{\alpha} - d_{\alpha}}{d_{\alpha}^2} \right) \right)
\]
(A.54)

where \( x_i = c_{\alpha_i} \) for \( \alpha_i, \ x_i = b_i \) for \( \alpha_2 \), and \( x_i = a \) for \( \alpha_3 \). Making use of the expectations in Lemma 9, we find
\[
E(\hat{\alpha}_l - \alpha_l) = E \left[ \frac{1}{T} \sum_{t=1}^{T} \left( \frac{\hat{c}_n}{\hat{d}_n} - \frac{c_n}{d_n} \right) \right] \\
= \frac{1}{T} \sum_{t=1}^{T} E(\hat{c}_n - c_n) \frac{1}{d_n} - \frac{1}{T} \sum_{t=1}^{T} E \left[ \left( \frac{\hat{c}_n}{\hat{d}_n} - \frac{c_n}{d_n} \right) \left( \hat{d}_n - d_n \right) \right] + \frac{1}{T} \sum_{t=1}^{T} \frac{E(\hat{d}_n - d_n)^2}{d_n^3} - \frac{E(\hat{d}_n - d_n)}{d_n^2}
\]

\[
= \frac{1}{T} \sum_{t=1}^{T} \left( \frac{(n+1)(c_n - 1)}{(T - L)d_n} + \frac{n}{Td_n} Z_{t-1}'A^{-1}Z_{t-1} \right) \\
- \frac{1}{T} \sum_{t=1}^{T} \left( \frac{2a(c_n - 1)^2 - 2b^2(c_n - 2)}{(T - L)d_n^2} + \frac{4(d_n - a)}{Td_n} Z_{t-1}'A^{-1}Z_{t-1} \right) \\
+ \frac{1}{T} \sum_{t=1}^{T} \frac{2a^2 + (n - 4)ad_n + (3 - 2n)d_n^2}{(T - L)d_n^3} - \frac{4a^2 - (n - 5)ad_n}{Td_n^3} Z_{t-1}'A^{-1}Z_{t-1} \right)
\]

\[
= \frac{1}{T} \sum_{t=1}^{T} \frac{(n+1)(c_n - 1)d_n^2 - 2a(c_n - 1)^2 d_n + 2b^2(c_n - 2)d_n + 2a^2c_n + (n - 4)ac_n d_n + (3 - 2n)c_n d_n^2}{(T - L)d_n^3}
\]

\[
+ \frac{1}{T} \sum_{t=1}^{T} \frac{n d_n^2 - 4(d_n - a)d_n + 4a^2c_n - (n - 5)ac_n d_n}{Td_n^3} Z_{t-1}'A^{-1}Z_{t-1} \right)
\]

\[
= \frac{1}{T} \sum_{t=1}^{T} \frac{2a^2c_n + \left[ -2a(c_n - 1)^2 + 2b^2(c_n - 2) + (n - 4)ac_n \right]d_n + \left[ (3 - 2n)c_n + (n + 1)(c_n - 1) \right]d_n^2}{(T - L)d_n^3}
\]

\[
+ \frac{1}{T} \sum_{t=1}^{T} \frac{-4a^2c_n + \left[ 4a - (n - 5)ac_n \right]d_n + n d_n^2 - 4d_n^2}{Td_n^3} Z_{t-1}'A^{-1}Z_{t-1} \right)
\]

(A.55)

continuing, we find
Next, we have
\[
E(\hat{\alpha}_2 - \alpha_2) = E\left[\frac{1}{T} \sum_{t=1}^{T} \left( \frac{\hat{b}_t}{d_n} - \frac{b_t}{d_n} \right) \right]
\]

\[
\approx \frac{1}{T} \sum_{t=1}^{T} \frac{E(\hat{b}_t - b_t)}{d_n} - \frac{1}{T} \sum_{t=1}^{T} E\left[\left( \frac{\hat{b}_t - b_t}{d_n} \right) \left( \frac{\hat{d}_n - d_n}{d_n^2} \right) \right] + \frac{1}{T} \sum_{t=1}^{T} b_t \left( \frac{E(\hat{d}_n - d_n)^2}{d_n^3} - \frac{E(\hat{d}_n - d_n)}{d_n^2} \right)
\]

\[
= \frac{1}{T(T-L)} \sum_{t=1}^{T} \left( \frac{n+1}{(T-L)d_n^2} b_t - \frac{2b_t d_n}{d_n^2} + \frac{2a^2 + (n-4)ad_n + (3-2n)d_n^2}{d_n^3} \right)
\]

\[
+ \frac{1}{T^2} \sum_{t=1}^{T} b_t \left( 4a^2 - (n-5)ad_n \right) Z'_{t-1} A^{-1} Z_{t-1}
\]

\[
= \frac{1}{T(T-L)} \sum_{t=1}^{T} \left( \frac{n+1}{d_n^2} b_t^2 - \frac{2b_t d_n}{d_n^2} + \frac{2a^2 + (n-4)ad_n + (3-2n)d_n^2}{d_n^3} \right)
\]

\[
- \frac{1}{T^2} \sum_{t=1}^{T} b_t \left( 4a^2 + (n-5)ad_n \right) Z'_{t-1} A^{-1} Z_{t-1}
\]

\[
= \frac{1}{T(T-L)} \sum_{t=1}^{T} \left( \frac{2a^2 + (n-4)ad_n - (n-2)d_n^2}{d_n^3} \right) - \frac{1}{T^2} \sum_{t=1}^{T} ab_t \left( 4a + (n-5)d_n \right) Z'_{t-1} A^{-1} Z_{t-1}
\]

\[
= \frac{1}{T} \sum_{t=1}^{T} \left( \frac{2a^2 + (n-4)ad_n - (n-2)d_n^2}{d_n^3(T-L)} - ab_t \frac{4a + (n-5)d_n}{d_n^3 T} \right) Z'_{t-1} A^{-1} Z_{t-1}
\]

(A.57)

and finally
\[
E(\hat{\alpha}_3 - \alpha_3) = E\left[ \frac{1}{T} \sum_{t=1}^{T} \left( \hat{a} - \frac{a}{d_n} \right) \right]
\]
\[
\equiv \frac{1}{T} \sum_{t=1}^{T} E\left( \hat{a} - \hat{\alpha} \right) d_n - \frac{1}{T} \sum_{t=1}^{T} \left[ E\left( \hat{a} - \hat{\alpha} \right) \left( \hat{d}_n - d_n \right) \right] + \frac{1}{T} \sum_{t=1}^{T} a \left( \frac{\hat{d}_n - d_n}{d_n^3} - \frac{\hat{d}_n - d_n}{d_n^2} \right)
\]
\[
\equiv \frac{1}{T} \sum_{t=1}^{T} \left( \frac{a(n+1)}{(T-L)d_n} \right) - \frac{1}{T} \sum_{t=1}^{T} \left( \frac{2ad_n}{(T-L)d_n^2} \right) + \frac{1}{T} \sum_{t=1}^{T} \left( \frac{2a^2 + (n-4)ad_n + (3-2n)d_n^2}{(T-L)d_n^2} \right)
\]
\[
\quad + \frac{1}{T} \sum_{t=1}^{T} a \left( \frac{4a^2 - (n-5)ad_n}{Td_n^3} \right) Z_{t-1}' A^{-1} Z_{t-1}
\]
\[
= \frac{1}{T} \sum_{t=1}^{T} \left( \frac{(n+1)d_n^2 - 2d_n^2 + 2a^2 + (n-4)ad_n + (3-2n)d_n^2}{(T-L)d_n^2} \right)
\]
\[
\quad + \frac{1}{T} \sum_{t=1}^{T} \left( \frac{a^2 - 4a + (n-5)d_n}{Td_n^3} \right) Z_{t-1}' A^{-1} Z_{t-1}
\]
\[
= \frac{a}{T} \sum_{t=1}^{T} \left( \frac{2a^2 + (n-4)ad_n - (n-2)d_n^2}{(T-L)d_n^3} - a \frac{4a + (n-5)d_n}{Td_n^3} \right) Z_{t-1}' A^{-1} Z_{t-1}
\]

(A.58)

Expectations of the form \( E\left( (\hat{\alpha}_i - \alpha_i)^2 \right) \) are as follows:

\[
E\left( (\hat{\alpha}_1 - \alpha_1)^2 \right) \equiv \frac{2}{T^2(T-L)} \sum_{k,j=1}^{T} \left( c_{ij} - 1 \right)^2 \frac{a\left(1-c_{ij}\right) + b_i^2c_n + b_j^2c_k + 2(c_{ij} - 1)d_{st}}{d_{ik}d_{kl}^2}
\]
\[
+ \frac{2}{T^2(T-L)} \sum_{k,j=1}^{T} \left( c_{ik}c_n - c_{kj}^2 \right) \frac{a^2(1+c_{ik}c_n - c_{kj}^2) + 2a(c_{ij} - 1)d_{st}}{d_{ik}d_{kl}^2}
\]
\[
+ \frac{4}{T^3} \sum_{k,j=1}^{T} \left( c_{ij} - 1 \right) \frac{2c_n - a}{d_{ik}d_{kl} - 2c_n \frac{d_{st}}{d_{ik}d_{kl}^2} + ac_{ik}c_n \frac{d_{st} - a}{d_{ik}d_{kl}^2}} Z_{k-1}' A^{-1} Z_{t-1}
\]
\[ E \left[ (\hat{\alpha}_2 - \alpha_2)^2 \right] \approx \frac{a}{T^3} \sum_{k,t=1}^{T} \frac{d_{kk}d_n}{d_{kk}^2d_n^2} + 4b_k b_t (d_{tt} - a) \frac{Z_{k-1}^T A^{-1} Z_{t-1}}{d_{kk}^2d_n^2} \]
\[ + \frac{1}{T^2 (T - L)} \sum_{k,t=1}^{T} \left( 2ab_k b_t (a + ac_{kk}c_{tt} - b_k b_t c_{tt}) + 2a b_t b_t (c_{tt} - 2) d_{tt} - 4b_k b_t^2 d_{kk}^2 + (ac_{kk} - a - 3b_k b_t) d_{kk} d_n \right) \frac{d_{kk}^2d_n^2}{d_{kk}^2d_n^2} \]

\[ E \left[ (\hat{\alpha}_3 - \alpha_3)^2 \right] \approx \frac{4a^3}{T^3} \sum_{k,t=1}^{T} \frac{d_{kk}^2d_n}{d_{kk}^2d_n^2} Z_{k-1}^T A^{-1} Z_{t-1}^{-1} \]
\[ + \frac{2a^2}{T^2 (T - L)} \sum_{k,t=1}^{T} \left( -ab_k b_t c_{tt} + a^2 + a^2 c_{kk} c_{tt} - b_k^2 d_{tt} - b_t^2 d_{kk} - 2a d_{tt} + a c_{kk} d_{tt} - d_{kk} d_n \right) \frac{d_{kk}^2d_n^2}{d_{kk}^2d_n^2} \] (A.59)

**Proof:** Begin by noting that each of these expectations has the form

\[ E \left[ \frac{1}{T} \sum_{i=1}^{T} \left( \frac{\hat{x}_i - x_i}{\hat{d}_{ii} - d_{ii}} \right)^2 \right] \]
\[ = \frac{1}{T^2} \sum_{k,t=1}^{T} \frac{x_k x_t}{d_{kk}^2d_n^2} E \left[ \left( \frac{\hat{x}_k - x_k}{\hat{d}_{kk}} \right)^2 \left( \frac{\hat{d}_{kk} - d_{kk}}{\hat{d}_{kk}} \right)^{-1} \left( \frac{\hat{x}_t - x_t}{\hat{d}_{tt}} \right)^2 \left( \frac{\hat{d}_{tt} - d_{tt}}{\hat{d}_{tt}} \right)^{-1} \right] \]
\[ \approx \frac{1}{T^2} \sum_{k,t=1}^{T} \frac{x_k x_t}{d_{kk}^2d_n^2} E \left[ \left( \frac{\hat{x}_k - x_k - \hat{d}_{kk} - d_{kk}}{\hat{d}_{kk} - d_{kk}} \right) \left( \frac{\hat{x}_t - x_t - \hat{d}_{tt} - d_{tt}}{\hat{d}_{tt} - d_{tt}} \right) \right] \]
\[ \approx \frac{1}{T^2} \sum_{k,t=1}^{T} \frac{x_k x_t}{d_{kk}^2d_n^2} \left( E \left[ (\hat{x}_k - x_k)(\hat{x}_t - x_t) \right] - E \left[ (\hat{x}_k - x_k)(\hat{d}_{tt} - d_{tt}) \right] \right) \]
\[ + \frac{1}{T^2} \sum_{k,t=1}^{T} \frac{x_k x_t}{d_{kk}^2d_n^2} \left( E \left[ (\hat{x}_t - x_t)(\hat{d}_{kk} - d_{kk}) \right] - E \left[ (\hat{x}_t - x_t)(\hat{d}_{tt} - d_{tt}) \right] \right) \]

(A.60)

where \( x_i = c_{ii} \) for \( \alpha_1 \), \( x_i = b_i \) for \( \alpha_2 \), and \( x_i = a \) for \( \alpha_3 \). Noting that the summation indices \((k,t)\) may be exchanged within an additive term of the summation, this may be written as
\[
E \left[ \frac{1}{T} \sum_{t=1}^{T} \left( \frac{\hat{x}_t}{d_{tt}} - \frac{x_t}{d_{tt}} \right) \right]^2
\]
\[
\approx \frac{1}{T^2} \sum_{k,t=1}^{T} \frac{x_k x_t}{d_{kk} d_{tt}} \left( E \left[ (\hat{x}_k - x_k) (\hat{x}_t - x_t) \right] \right) - \frac{2}{x_k d_{tt}} \left( E \left[ (\hat{x}_k - x_k) (\hat{d}_{tt} - d_{tt}) \right] \right) + \frac{1}{d_{kk} d_{tt}} \left( E \left[ (\hat{d}_{kk} - d_{kk}) (\hat{d}_{tt} - d_{tt}) \right] \right)
\]
\[
= \frac{1}{T^2} \sum_{k,t=1}^{T} \left( E \left[ (\hat{x}_k - x_k) (\hat{x}_t - x_t) \right] \right) - \frac{2 x_k E \left[ (\hat{x}_k - x_k) (\hat{d}_{tt} - d_{tt}) \right]}{d_{kk} d_{tt}^2} + \frac{x_k x_t E \left[ (\hat{d}_{kk} - d_{kk}) (\hat{d}_{tt} - d_{tt}) \right]}{d_{kk}^2 d_{tt}^2}
\]

(A.61)

We find, using the expectations from Lemma 9, as follows:
\[ E\left[ (\hat{\alpha}_1 - \alpha_1)^2 \right] \]
\[ \approx \frac{1}{T^2} \sum_{k,t=1}^{T} \left\{ \frac{E\left[ (\hat{c}_{k^*} - c_{k^*})(\hat{c}_{t^*} - c_{t^*}) \right]}{d_{kk} d_{tt}} - \frac{2c_n E\left[ (\hat{c}_{k^*} - c_{k^*})(\hat{d}_{t^*} - d_{t^*}) \right]}{d_{kk} d_{tt}^2} + c_{kk} c_n E\left[ (\hat{d}_{k^*} - d_{k^*})(\hat{d}_{t^*} - d_{t^*}) \right] \right\} \]
\[ \approx \frac{1}{T^2} \sum_{k,t=1}^{T} \left( \frac{2(c_{kk} - 1)^2}{T - L} + \frac{4(c_{kk} - 1)Z'_{k-1}A^{-1}Z_{t-1}}{T} \right) \]
\[ + \frac{1}{T^2} \sum_{k,t=1}^{T} \left( \frac{2c_n \left[ 2a(c_{kk} - 1)^2 - 4b_k b_t (c_{kk} - 1) + 2b_t^2 c_n + 4(d_{kk} - a)Z'_{k-1}A^{-1}Z_{t-1} \right]}{T - L} \right) \]
\[ + \frac{1}{T^2} \sum_{k,t=1}^{T} \left( \frac{c_{kk} c_n \left[ -2ab_k b_t c_{kk} + 2a^2 + 2a^2 c_{kk} c_n - 2b_t^2 d_{tt} - 2b_t^2 d_{kk} - 4a d_{tt} + 2a c_{kk} d_{tt} + 4 \frac{a d_{tt} - a^2}{T} Z'_{k-1}A^{-1}Z_{t-1} \right]}{d_{kk} d_{tt}^2} \right) \]
\[ \approx \frac{1}{T^2} \sum_{k,t=1}^{T} \left( \frac{2(c_{kk} - 1)^2}{(T - L)d_{kk} d_{tt} + \frac{4(c_{kk} - 1)Z'_{k-1}A^{-1}Z_{t-1}}{Td_{kk} d_{tt}} \right) \]
\[ + \frac{1}{T^2} \sum_{k,t=1}^{T} \left( -4c_n \frac{a(c_{kk} - 1)^2 - 2b_k b_t (c_{kk} - 1) + b_t^2 c_n}{(T - L)d_{kk} d_{tt}^2} - \frac{8c_n (d_{kk} - a)}{Td_{kk} d_{tt}^2} Z'_{k-1}A^{-1}Z_{t-1} \right) \]
\[ + \frac{1}{T^2} \sum_{k,t=1}^{T} \left( 2c_{kk} c_n \frac{-ab_k b_t c_{kk} + a^2 + a^2 c_{kk} c_n - b_t^2 d_{tt} - b_t^2 d_{kk} - 2a d_{tt} + a c_{kk} d_{tt} + 4c_{kk} c_n \frac{a d_{tt} - a^2}{T} Z'_{k-1}A^{-1}Z_{t-1} }{(T - L)d_{kk} d_{tt}^2} \right) \]

(A.62)

continuing, we find
\[
E\left[ (\hat{\alpha}_t - \alpha_t)^2 \right] \approx \frac{1}{T^2} \sum_{k,t=1}^{T} \left( \frac{2(c_{kt} - 1)^2}{(T - L)d_{kk}d_{tt}} - 4c_{tt} \frac{a(c_{kt} - 1)^2 - 2b_{kt}^2(c_{kt} - 1) + b_{kt}^2c_{tt}}{(T - L)d_{kk}d_{tt}} \right) \\
+ \frac{1}{T^2} \sum_{k,t=1}^{T} \left( 2c_{kk}c_{tt} - ab_{kt}b_{kt} + a^2 + a^2 c_{kk}c_{tt} - b_{kt}^2d_{tt} - b_{kt}^2d_{kk} - 2a d_{tt} + a c_{tt}d_{tt} \right) \left( T - L \right) d_{kk}^2 d_{tt}^2 \\
+ \frac{1}{T^2} \sum_{k,t=1}^{T} \left( \frac{4(c_{kt} - 1) - 8c_{tt}(d_{tt} - a)}{Td_{kk}d_{tt}} + 4c_{kk}c_{tt} \frac{a d_{tt} - a^2}{Td_{kk}^2 d_{tt}^2} \right) Z_{k-1} A^{-1} Z_{t-1}
\]

(A.63)

Making use of the fact that

\[
a(c_{kt} - 1)^2 - 2b_{kt}b_{kt}(c_{kt} - 1) + b_{kt}^2c_{tt} = ac_{kt}^2 - 2ac_{kt} + a - 2b_{kt}b_{kt}c_{kt} + 2b_{kt}b_{kt} + b_{kt}^2c_{tt} \\
= (ac_{kt}^2 + 2ac_{kt}^2) + a - 2b_{kt}b_{kt}c_{kt} + b_{kt}^2c_{tt} - 2(ac_{kt} - b_{kt}b_{kt}) \\
= -ac_{kt}^2 + a + b_{kt}^2c_{tt} + 2c_{kt}(ac_{kt} - b_{kt}b_{kt}) - 2d_{tt} \\
= a - ac_{kt}^2 + b_{kt}^2c_{tt} + 2c_{kt}d_{tt} - 2d_{tt} \\
= a - ac_{kt}^2 + b_{kt}^2c_{tt} + 2(c_{kt} - 1)d_{tt}
\]

and also replacing \(-b_{kt}^2d_{tt} - b_{kt}^2d_{kk}\) with \(-2b_{kt}^2d_{kk}\) (which is permissible because its multipliers are exchangeable in \((k,t)\)) we find
\[
E \left[ (\hat{\alpha}_t - \alpha_t)^2 \right] \approx \frac{1}{T^2} \sum_{k,t=1}^{T} \left\{ \frac{2(c_{kt} - 1)^2}{(T-L)d_{kk}d_n} - 4c_n a - ac_{kt}^2 + b_k^2 c_n + 2(c_{kt} - 1)d_{kk} \right\} \\
+ \frac{1}{T^2} \sum_{k,t=1}^{T} \left\{ 2c_{tt} c_n - ab_k b_k c_{tt} + a^2 + a^2 c_{kk} c_n - 2b_k^2 d_{kk} - 2a d_{tt} + a c_k d_{tt} \right\} \\
+ \frac{1}{T^2} \sum_{k,t=1}^{T} \left\{ \frac{4(c_{tt} - 1)}{Td_{kk}d_n} - 8c_n (d_{tt} - a) + 4c_{kk} c_n \frac{ad_{tt} - a^2}{Td_{kk}^2 d_n^2} \right\} Z'_{k-1} A^{-1} Z_{t-1}
\]

\[
E \left[ (\hat{\beta}_t - \beta_t)^2 \right] \approx \frac{1}{T^2} \sum_{k,t=1}^{T} \left\{ \frac{2(c_{kt} - 1)^2}{(T-L)d_{kk}d_n} - 4c_n a - ac_{kt}^2 + b_k^2 c_n + 2(c_{kt} - 1)d_{kk} \right\} \\
+ \frac{1}{T^2} \sum_{k,t=1}^{T} \left\{ 2c_{kk} c_n - ab_k b_k c_{tt} + a^2 + a^2 c_{kk} c_n + a(c_{kt} - 2)d_{kk} + 2c_{kk} c_n \frac{-2b_k^2}{(T-L)d_{kk}^2 d_n^2} \right\} \\
+ \frac{1}{T^2} \sum_{k,t=1}^{T} \left\{ \frac{4(c_{tt} - 1)}{Td_{kk}d_n} - 8c_n (d_{tt} - a) + 4c_{kk} c_n \frac{ad_{tt} - a^2}{Td_{kk}^2 d_n^2} \right\} Z'_{k-1} A^{-1} Z_{t-1}
\]

\[= \frac{1}{T^2} \sum_{k,t=1}^{T} \left\{ \frac{2(c_{tt} - 1)^2}{(T-L)d_{kk}d_n} - 4c_n a(1 - c_{tt}^2) + b_k^2 c_n + b_k^2 c_{kk} + 2(c_{tt} - 1)d_{kk} \right\} \\
+ \frac{1}{T^2} \sum_{k,t=1}^{T} \left\{ 2c_{kk} c_n - ab_k b_k c_{tt} + a^2 + a^2 c_{kk} c_n + a(c_{tt} - 2)d_{kk} \right\} \\
+ \frac{1}{T^2} \sum_{k,t=1}^{T} \left\{ \frac{4(c_{tt} - 1)}{Td_{kk}d_n} - 8c_n (d_{tt} - a) + 4c_{kk} c_n \frac{ad_{tt} - a^2}{Td_{kk}^2 d_n^2} \right\} Z'_{k-1} A^{-1} Z_{t-1}
\]

We next use
\(-ab_k b_t c_{kt} = a^2 c_{kt}^2 - ab_k b_t c_{kt} - a^2 c_{kt}^3 = ac_{kt} (ac_{kt} - b_k b_t) - a^2 c_{kt}^3 = ac_{kt} d_{ks} - a^2 c_{kt}^3\)

to find

\[
E\left[ (\hat{\alpha}_t - \alpha_t)^2 \right] \approx \frac{1}{T^2} \sum_{k,t=1}^{T} \left( \frac{2(c_{kt} - 1)^2}{(T-L)d_{kt}d_{nt}} - 4c_{nt} a\left(1-c_{kt}^2\right) + b_k^2 c_{nt} + b_t^2 c_{nk} + 2(c_{kt} - 1)d_{ks} \right) \\
+ \frac{1}{T^2} \sum_{k,t=1}^{T} \left( \frac{2c_{ik} c_{nt} d_{kt} - a^2 c_{kt}^2 + a^2 + a^2 c_{ik} c_{nt} + a(c_{kt} - 2)d_{kt}}{(T-L)d_{ik}^2 d_{nt}^2} \right) \\
+ \frac{1}{T^2} \sum_{k,t=1}^{T} \left( \frac{4(c_{kt} - 1) - 8c_{nt} (d_{kt} - a)}{Td_{nt} d_{nt}^2} + 4c_{ik} c_{nt} \frac{ad_{kt} - a^2}{Td_{nt} d_{nt}^2} \right) Z'_{k-1} A^{-1} Z_{t-1} \\
= \frac{2}{T^2(T-L)} \sum_{k,t=1}^{T} \left( \frac{(c_{kt} - 1)^2}{d_{kt} d_{nt}} - 2c_{nt} a\left(1-c_{kt}^2\right) + b_k^2 c_{nt} + b_t^2 c_{nk} + 2(c_{kt} - 1)d_{ks} \right) \\
+ \frac{2}{T^2(T-L)} \sum_{k,t=1}^{T} \left( \frac{a^2 (1+c_{ik} c_{nt} - c_{kt}^2) + 2a(c_{kt} - 1)d_{st}}{d_{ik}^2 d_{nt}^2} \right) \\
+ \frac{4}{T^2} \sum_{k,t=1}^{T} \left( \frac{c_{ks} - 1}{d_{ik} d_{nt}} - 2c_{nt} \frac{d_{kt} - a}{d_{nt} d_{nt}^2} + ac_{ik} c_{nt} \frac{d_{kt} - a}{d_{nt}^2 d_{nt}^2} \right) Z'_{k-1} A^{-1} Z_{t-1}
\]

(A.67)

Next is
\[
E\left[ (\hat{\alpha}_2 - \alpha_2)^2 \right] \\
\approx \frac{1}{T^2} \sum_{k,t=1}^{T} \left( \frac{E\left[ (\hat{b}_k - b_k)(\hat{b}_t - b_t) \right]}{d_{kk}d_{tt}} - \frac{2b_tE\left[ (\hat{b}_k - b_k)(\hat{d}_n - d_n) \right]}{d_{kk}d_{nn}^2} + \frac{b_kb_tE\left[ (\hat{d}_{kk} - d_{kk})(\hat{d}_n - d_n) \right]}{d_{nn}^2} \right) \\
\approx \frac{1}{T^2} \sum_{k,t=1}^{T} \left( \frac{a(c_{kt} - 1) + b_kb_t}{T - L} + \frac{a}{T} Z'_{k-1}A^{-1}Z_{t-1} \right) + \frac{1}{T^2} \sum_{k,t=1}^{T} \left( -\frac{2b_t}{T - L} \right) \\
+ \frac{1}{T^2} \sum_{k,t=1}^{T} \left( \frac{b_kb_t}{T - L} \left[ -2ab_k b_t c_{kt} + 2a^2 + 2a^2 c_{kk} c_{tt} - 2b_k^2 d_{tt} - 2b_t^2 d_{kk} - 4a d_{tt} + 2a c_{kt} d_{tt} + 4 \frac{ad_{tt} - a^2}{T} Z'_{k-1}A^{-1}Z_{t-1} \right] \right) \\
= \frac{1}{T^2} \sum_{k,t=1}^{T} \left( \frac{a(c_{kt} - 1) + b_kb_t}{T - L} d_{kk}d_{tt} - \frac{4b_kb_td_{tt}}{(T - L)d_{kk}d_{tt}} + \frac{a}{Td_{kk}d_{tt}} Z'_{k-1}A^{-1}Z_{t-1} + 4b_kb_t \frac{ad_{tt} - a^2}{Td_{kk}d_{tt}} Z'_{k-1}A^{-1}Z_{t-1} \right) \\
+ \frac{1}{T^2} \sum_{k,t=1}^{T} \left( \frac{b_kb_t}{T - L} \left[ -2ab_k b_t c_{kt} + 2a^2 + 2a^2 c_{kk} c_{tt} - 2b_k^2 d_{tt} - 2b_t^2 d_{kk} - 4a d_{tt} + 2a c_{kt} d_{tt} \right] \right) \\
= \frac{1}{T^2} \sum_{k,t=1}^{T} \left( \frac{a(c_{kt} - 1) - 3b_kb_t d_{kk}}{(T - L)d_{kk}^2d_{tt}^2} \right) \\
+ \frac{1}{T^2} \sum_{k,t=1}^{T} \left( \frac{-2ab_k b_t c_{kt} + 2a^2 b_t b_t + 2a^2 b_k b_k c_{tt} - 2b_k^2 b_t d_{tt} - 2b_t^2 b_k d_{kk} - 4a b_t b_t d_{tt} + 2a b_t b_k c_{kt} d_{tt}}{(T - L)d_{kk}^2d_{tt}^2} \right) \\
+ \frac{1}{T^2} \sum_{k,t=1}^{T} \left( \frac{ad_{tt} + 4ab_k b_t (d_{tt} - a)}{Td_{kk}d_{tt}^2} \right) \right) Z'_{k-1}A^{-1}Z_{t-1} \\
= \frac{1}{T^2} \sum_{k,t=1}^{T} \left( \frac{-2ab_k^2 b_t c_{kt} + 2a^2 b_k b_t + 2a^2 b_k b_k c_{tt} - 4a b_t b_t d_{tt} + 2a b_t b_k c_{tt} d_{tt} - 2b_k^2 b_k d_{tt} - 2b_k^2 b_k d_{kk}}{(T - L)d_{kk}^2d_{tt}^2} \right) \\
+ \frac{1}{T^2} \sum_{k,t=1}^{T} \left( \frac{ac_{tt} - a - 3b_kb_t}{(T - L)d_{kk}d_{tt}^2} \right) + \frac{1}{T^2} \sum_{k,t=1}^{T} \left( \frac{ad_{tt} + 4ab_k b_t (d_{tt} - a)}{Td_{kk}d_{tt}^2} \right) \right) Z'_{k-1}A^{-1}Z_{t-1}
\]
exchanging summation indices to replace \( b_k^3 b_t d_n \) with \( b_k^3 b_t d_{kk} \), we find

\[
E\left[ (\alpha_2 - \alpha_2)^2 \right] \approx \frac{a}{T^3} \sum_{k,t=1}^{T} \frac{d_{kk} d_{tt} + 4 b_k b_t (d_{tt} - a)}{d_{kk}^2 d_{tt}^2} Z_{k-1} Z_{t-1} \]

\[
+ \frac{1}{T^2 (T - L)} \sum_{k,t=1}^{T} \left( \frac{2 a b_k b_t (a + ac_k c_t - b_k^2 c_{kk}) + 2 a b_k b_t (c_{tt} - 2) d_{tt} - 4 b_k b_t d_{kk} + (ac_{tt} - a - 3 b_k b_t) d_{kk} d_{tt}}{d_{kk}^2 d_{tt}^2} \right)
\]

(A.69)

Next is
\[
E\left[ (\hat{\alpha}_i - \alpha_i)^2 \right]
\]
\[
\approx \frac{1}{T^2} \sum_{k,t=1}^{T} \left( \frac{E\left[ (\hat{\alpha} - a)^2 \right]}{d_{kk}^2 d_{tt}^2} - \frac{2a E\left[ (\hat{\alpha} - a)(\hat{\alpha} - d) \right]}{d_{kk}^2 d_{tt}^2} + \frac{a^2 E\left[ (\hat{\alpha}_k - d_k)(\hat{\alpha}_t - d_t) \right]}{d_{kk}^2 d_{tt}^2} \right)
\]
\[
\approx \frac{1}{T^2} \sum_{k,t=1}^{T} \left( \frac{2a^2}{T - L} \frac{T}{d_{kk} d_{tt}} \right) + \frac{1}{T^2} \sum_{k,t=1}^{T} \left( -\frac{2a^2}{(T - L) d_{kk} d_{tt}} \right)
\]
\[
+ \frac{1}{T^2} \sum_{k,t=1}^{T} \left( \frac{a^2 (-2ab k b c_{kt} + 2a^2 c_{kk} c_{tt} - 2b^2 d_{tt} - 2b^2 d_{kk} - 4a d_{kt} + 2a c_{kt} d_{kt} + 4 a d_{kt} - a^2 Z'_{k-1} A^{-1} Z_{k-1})}{d_{kk}^2 d_{tt}^2} \right)
\]
\[
= \frac{1}{T^2} \sum_{k,t=1}^{T} \left( \frac{2a^2}{T - L} \frac{T}{d_{kk}^2 d_{tt}^2} \right) + \frac{1}{T^2} \sum_{k,t=1}^{T} \left( \frac{4a^2 d_{kt} - 4a^4 d_{kk}^2}{T d_{kk}^2 d_{tt}^2} Z'_{k-1} A^{-1} Z_{k-1} \right)
\]
\[
+ \frac{1}{T^2} \sum_{k,t=1}^{T} \left( \frac{a^2 (-2ab k b c_{kt} + 2a^2 c_{kk} c_{tt} - 2b^2 d_{tt} - 2b^2 d_{kk} - 4a d_{kt} + 2a c_{kt} d_{kt})}{(T - L) d_{kk}^2 d_{tt}^2} \right)
\]
\[
= \frac{4a^3}{T^3} \sum_{k,t=1}^{T} \frac{d_{kt} - a}{d_{kk}^2 d_{tt}^2} Z'_{k-1} A^{-1} Z_{k-1}
\]
\[
+ \frac{2a^2}{T^2 (T - L)} \sum_{k,t=1}^{T} \frac{-ab k b c_{kt} + a^2 + a^2 c_{kk} c_{tt} - b^2 d_{tt} - b^2 d_{kk} - 2a d_{kt} + a c_{kt} d_{kt} - d_{kt} d_{tt}}{d_{kk}^2 d_{tt}^2}
\]

(A.70)

Expectations of the form \( E\left[ (\hat{\alpha}_i - \alpha_i)(\hat{\alpha}_j - \alpha_j) \right] \) with \( i \neq j \) are as estimated as follows:
\[
E[(\hat{\alpha}_1 - \alpha_1)(\hat{\alpha}_2 - \alpha_2)] \approx -\frac{2\alpha_2(\alpha_3 + a\alpha_1)}{(T - L)a} + \frac{2}{T^3} \sum_{k,l=1}^{T} \left( \frac{b_k}{d_{kl}d_n} - \frac{2b_k(d_{kl} - a)}{d_{kl}^2 d_n^2} + \frac{2ab_k c_n (d_{kl} - a)}{d_{kl}^2 d_n^2} \right) Z_{k-1}' A^{-1} Z_{l-1} \\
+ \frac{2}{T^2(T - L)} \sum_{k,l=1}^{T} \left( \frac{b_k}{d_{kl}d_n} - \frac{a}{d_{kl}^2} \left( b_k c_{kl} - b_l \right) + b_k \left( b_k + b_l \right) c_n + b_k^2 c_{kl} \right) \\
+ \frac{2}{T^2(T - L)} \sum_{k,l=1}^{T} \left( b_k c_n - \frac{a}{d_{kl}^2} b_k c_{kl} + a^2 c_{kl} c_n - 2a d_{kl} + a c_{kl} d_{kl} \right)
\]

\[
E[(\hat{\alpha}_1 - \alpha_1)(\hat{\alpha}_3 - \alpha_3)] \approx -\frac{2\alpha_3}{(T - L)a} + \frac{2a}{T^2(T - L)} \sum_{k,l=1}^{T} \left( a (c_{kl} - 1)^2 - 2b_k b_l (c_{kl} - 1) + b_k^2 c_n + b_l^2 (c_{kl} + c_n) \right) \\
+ \frac{2}{T^2(T - L)} \sum_{k,l=1}^{T} \left( a^2 c_n - \frac{b_k b_l c_{kl} + a + a c_{kl} c_n + (c_{kl} - 2)d_{kl}}{d_{kl}^2 d_n^2} \right) - \frac{4a}{T^3} \sum_{k,l=1}^{T} \left( a - d_{kl} \right) b_k^2 Z_{k-1}' A^{-1} Z_{l-1}
\]

and

\[
E[(\hat{\alpha}_2 - \alpha_2)(\hat{\alpha}_3 - \alpha_3)] \approx -\frac{2\alpha_2 \alpha_3}{(T - L)} + \frac{4a^2}{T^3} \sum_{k,l=1}^{T} \left( b_l (d_{kl} - a) \right) Z_{k-1}' A^{-1} Z_{l-1} \\
+ \frac{2a}{T^2(T - L)} \sum_{k,l=1}^{T} \left( a c_{kl} c_n - b_k b_l c_{kl} + (c_{kl} - 2)d_{kl} \right) - \frac{b_l^2 (b_k + b_l)}{d_{kl}^2 d_n^2}
\]

(A.71)

To see this note that each of these expectations has the form
where \( x_i \) or \( y_i \) is \( c_n \) for \( \alpha_1 \), \( x_i \) or \( y_i \) is \( b_i \) for \( \alpha_2 \), and \( x_i \) or \( y_i \) is \( a \) for \( \alpha_3 \), and we exchanged the summation indices in the final summation. Continuing, we have

\[
E\left\{ \left[ \frac{1}{T} \sum_{t=1}^{T} \left( \frac{x_t}{d_n} - \frac{x_i}{d_a} \right) \right] \left[ \frac{1}{T} \sum_{t=1}^{T} \left( \frac{y_t}{d_n} - \frac{y_i}{d_a} \right) \right] \right\}
\]

\[
= \frac{1}{T^2} \sum_{k,t=1}^{T} \frac{x_k y_t}{d_{kk} d_n} \sum_{i=1}^{T} \left[ \frac{1 + (\hat{x}_k - x_k) / x_k}{1 + (\hat{d}_{kk} - d_{kk}) / d_{kk}} \right] \left[ \frac{1 + (\hat{y}_i - y_i) / y_i}{1 + (\hat{d}_n - d_n) / d_n} \right]
\]

\[
= \frac{1}{T^2} \sum_{k,t=1}^{T} \frac{x_k y_t}{d_{kk} d_n} \left[ \frac{(\hat{x}_k - x_k) / x_k}{\hat{d}_{kk} - d_{kk}} \right] \left[ \frac{\hat{y}_i - y_i}{y_i - \hat{d}_n / d_n} \right] \left[ \frac{\hat{d}_n - d_n}{d_n} \right]
\]

\[
= \frac{1}{T^2} \sum_{k,t=1}^{T} \frac{x_k y_t}{d_{kk} d_n} \left[ \frac{E(\hat{x}_k - x_k)(\hat{y}_i - y_i)}{x_k y_t} \right] \left[ \frac{E(\hat{d}_{kk} - d_{kk})(\hat{d}_n - d_n)}{d_{kk} d_n} \right] + \frac{1}{T^2} \sum_{k,t=1}^{T} \frac{y_k x_t}{d_{kk} d_n} \left[ \frac{E(\hat{d}_{kk} - d_{kk})(\hat{d}_n - d_n)}{d_{kk} d_n} \right]
\]

(E.72)

Evaluating these expectations we find
\[ E \left[ (\hat{\alpha}_1 - \alpha_1)(\hat{\alpha}_2 - \alpha_2) \right] \equiv \frac{1}{T^2} \sum_{k,j=1}^{T} \left( \frac{E \left[ (\hat{\beta}_j - \beta_j)(\hat{c}_{kk} - c_{kk}) \right]}{d_{kk}d_{nt}} - \frac{E \left[ (\hat{c}_{kk} - c_{kk})(\hat{d}_n - d_n) \right]}{d_{kk}^2d_{nt}^2} \right) \]

\[ + \frac{1}{T^2} \sum_{k,j=1}^{T} \left( \frac{-c_n}{d_{kk}d_{nt}^2} \right) \]

\[ + \frac{1}{T^2} \sum_{k,j=1}^{T} \left( \frac{2b_k(c_{st} - 1)}{T - L} + \frac{2b_k}{T} Z'_{k-1}A^{-1}Z_{t-1} \right) \]

\[ + \frac{1}{T^2} \sum_{k,j=1}^{T} \left( \frac{-2a(c_{st} - 1)^2 - 4b_kb_k(c_{st} - 1) + 2b^2c_n^2}{T - L} - \frac{4(d_{st} - a)}{T} Z'_{k-1}A^{-1}Z_{t-1} \right) \]

\[ + \frac{1}{T^2} \sum_{k,j=1}^{T} \left( \frac{2b_kd_n}{T} \right) \]

\[ + \frac{1}{T^2} \sum_{k,j=1}^{T} \left( \frac{-c_n}{d_{kk}d_{nt}^2} \right) \]

\[ + \frac{1}{T^2} \sum_{k,j=1}^{T} \left( \frac{-2ab_k b_k c_{st} + 2a^2 + 2a^2c_{kk}c_n - 2b^2d_n - 2b^2d_k - 2b^2d_n - 4a d_k + 2a c_n d_k + 4a d_k - a^2}{T} \right) \]

\[ + \frac{1}{T^2} \sum_{k,j=1}^{T} \left( \frac{b_n c_n}{T - L} \right) \]

\[ = \frac{1}{T^2} \sum_{k,j=1}^{T} \left( \frac{2b_k}{T} \right) \]

\[ + \frac{1}{T^2} \sum_{k,j=1}^{T} \left( \frac{-2b_k b_k c_{st} + 2a^2 + 2a^2c_{kk}c_n - 2b^2d_n - 2b^2d_k - 2b^2d_n - 4a d_k + 2a c_n d_k + 4a d_k - a^2}{T} \right) \]

\[ + \frac{1}{T^2} \sum_{k,j=1}^{T} \left( \frac{b_n c_n}{T - L} \right) \]

\[ = \frac{2}{T^2} \sum_{k,j=1}^{T} \left( \frac{b_k}{d_{kk}d_{nt}^2} \right) - \frac{2b_k b_k c_{st} + 2a^2 + 2a^2c_{kk}c_n - 2b^2d_n - 2b^2d_k - 2b^2d_n - 4a d_k + 2a c_n d_k + 4a d_k - a^2}{T} \right) \]

\[ + \frac{2}{T^2} \sum_{k,j=1}^{T} \left( \frac{b_n c_n}{d_{kk}d_{nt}^2} \right) \]

\[ + \frac{2}{T^2} \sum_{k,j=1}^{T} \left( \frac{b_k c_n}{T - L} \right) \]

\[ + \frac{2}{T^2} \sum_{k,j=1}^{T} \left( \frac{b_k c_n}{T - L} \right) \]
Manipulating the final summation (including an exchange of the indices) we have

\[
\frac{2}{T^2} \sum_{k,t=1}^{T} \left( b_k c_{k,t} \frac{-ab_k b_t c_{k,t} + a^2 + a^2 c_{k,k} c_{t,t} - b_k^2 d_{k,t} - b_t^2 d_{t,t} - 2a d_{k,t} + a c_{k,t} d_{t,t}}{(T - L) d_{k,k}^2 d_{t,t}^2} \right)
\]

which leads to

\[
\frac{2}{T^2} \sum_{k,t=1}^{T} \left( b_k c_{k,t} \frac{-ab_k b_t c_{k,t} + a^2 + a^2 c_{k,k} c_{t,t} - 2a d_{k,t} + a c_{k,t} d_{t,t} + b_k c_{k,t} - b_t^2 d_{t,t} - b_k^2 d_{k,k}}{(T - L) d_{k,k}^2 d_{t,t}^2} \right)
\]

\[
= \frac{2}{T^2} \sum_{k,t=1}^{T} \left( b_k c_{k,t} \frac{-ab_k b_t c_{k,t} + a^2 + a^2 c_{k,k} c_{t,t} - 2a d_{k,t} + a c_{k,t} d_{t,t} + b_k c_{k,t} - b_t^2 d_{t,t} - b_k^2 c_{k,t} d_{k,k}}{(T - L) d_{k,k}^2 d_{t,t}^2} \right)
\]

\[
\frac{2}{T^2} \sum_{k,t=1}^{T} \left( b_k c_{k,t} \frac{-ab_k b_t c_{k,t} + a^2 + a^2 c_{k,k} c_{t,t} - 2a d_{k,t} + a c_{k,t} d_{t,t}}{(T - L) d_{k,k}^2 d_{t,t}^2} \right)
\]

which leads to
\[
E[(\hat{\alpha}_1 - \alpha_1)(\hat{\alpha}_2 - \alpha_2)] \approx \frac{2}{T^3} \sum_{k,t=1}^T \left( \frac{b_k}{d_{kk} d_{tt}} - \frac{2b_t(d_{kt} - a)}{d_{kk} d_{tt}^2} + \frac{2ab_k c_n(d_{kt} - a)}{d_{kk}^2 d_{tt}^2} \right) Z_{t-1}' A^{-1} Z_{t-1} \\
+ \frac{2}{T^2} \sum_{k,t=1}^T \left( \frac{b_k (c_{kt} - 1) - b_k c_n}{(T - L)d_{kk} d_{tt}} - b_t a (c_{kt} - 1)^2 - 2b_t b_k (c_{kt} - 1) + b_k^2 c_n \frac{1}{(T - L)d_{kk} d_{tt}^2} \right) \\
+ \frac{2}{T^2} \sum_{k,t=1}^T \left( \frac{b_k c_n - ab_k b_k c_n + a^2}{(T - L)d_{kk} d_{tt}^2} - b_t a (c_{kt} - 1)^2 - 2b_t b_k (c_{kt} - 1) + b_k^2 c_n \right) \\
= \frac{2}{T^3} \sum_{k,t=1}^T \left( \frac{b_k}{d_{kk} d_{tt}} - \frac{2b_t(d_{kt} - a)}{d_{kk} d_{tt}^2} + \frac{2ab_k c_n(d_{kt} - a)}{d_{kk}^2 d_{tt}^2} \right) Z_{t-1}' A^{-1} Z_{t-1} \\
+ \frac{2}{T^2} \sum_{k,t=1}^T \left( \frac{b_k (c_{kt} - 1) - b_k c_n}{(T - L)d_{kk} d_{tt}} - b_t a (c_{kt} - 1)^2 - 2b_t b_k (c_{kt} - 1) + b_k^2 c_n \frac{1}{(T - L)d_{kk} d_{tt}^2} \right) \\
+ \frac{2}{T^2} \sum_{k,t=1}^T \left( \frac{b_k c_n - ab_k b_k c_n + a^2}{(T - L)d_{kk} d_{tt}^2} - b_t a (c_{kt} - 1)^2 - 2b_t b_k (c_{kt} - 1) + b_k^2 c_n \right) \\
\]

We next make use of the facts that \(\Sigma(c_n / d_n) = T\alpha_1\), and \(\Sigma b_t / d_n = T\alpha_2\), and \(\Sigma(1 / d_n) = T\alpha_3 / a\) so that

\[
\frac{2}{T^2} \sum_{k,t=1}^T \frac{-b_k - b_k c_n}{(T - L)d_{kk} d_{tt}} = - \frac{2}{T^2} \sum_{k,t=1}^T \frac{b_k}{d_{kk} d_{tt}} - \frac{2}{T^2} \sum_{k,t=1}^T \frac{b_k c_n}{d_{kk} d_{tt}} \\
= - \frac{2}{T^2} \left( \frac{T}{T - L} \sum_{k=1}^T b_k \right) \left( \frac{1}{d_{kk}} \sum_{t=1}^T d_{tt} \right) - \frac{2}{T^2} \left( \frac{T}{T - L} \sum_{k=1}^T b_k c_n \right) \left( \frac{1}{d_{kk}} \sum_{t=1}^T d_{tt} \right) \\
= - \frac{2 \alpha_3 \alpha_3}{(T - L) a} - \frac{2 \alpha_3 \alpha_2}{(T - L)} = - \frac{2}{(T - L)} \left( \alpha_3 \alpha_3 + \alpha_3 \alpha_2 \right) = - \frac{2 \alpha_2 (\alpha_3 + \alpha_3)}{a(T - L)} \\
\]

and therefore
\[ E[(\hat{\alpha}_1 - \alpha_1)(\hat{\alpha}_2 - \alpha_2)] \approx \frac{2}{T^2} \sum_{k,t=1}^{T} \left( \frac{-b_k - b_t c_n}{(T-L)d_{kk}d_{nt}} \right) \]
\[
+ \frac{2}{T^3} \sum_{k,t=1}^{T} \left( \frac{b_k d_{kk}}{d_{kk}d_{nt}^2} - \frac{2b_k (d_{kt} - a)}{d_{kk}d_{nt}^2} + \frac{2ab_k c_n (d_{kt} - a)}{d_{kk}^2 d_{nt}^2} \right) Z_{k-1} A^{-1} Z_{t-1} \]
\[
+ \frac{2}{T^2} \sum_{k,t=1}^{T} \left( \frac{b_k c_{kt}}{(T-L)d_{kk}d_{nt}} - b_t \frac{a(c_{ks} - 1)^2 - 2b_k b_t (c_{ks} - 1) + b_k (b_k + b_t) c_{nt} + b_t^2 c_{kk}}{(T-L)d_{kk}d_{nt}^2} \right) \]
\[
+ \frac{2}{T^2} \sum_{k,t=1}^{T} \left( \frac{b_k c_{nt}}{(T-L)d_{kk}d_{nt}} - b_t \frac{a(c_{ks} - 1)^2 - 2b_k b_t (c_{ks} - 1) + b_k (b_k + b_t) c_{nt} + b_t^2 c_{kk}}{(T-L)d_{kk}d_{nt}^2} \right) \]
\[
= -\frac{2\alpha_2 \alpha_3 (\alpha_3 + \alpha a_1)}{(T-L)a} + \frac{2}{T^3} \sum_{k,t=1}^{T} \left( \frac{b_k d_{kk}}{d_{kk}d_{nt}^2} - \frac{2b_k (d_{kt} - a)}{d_{kk}d_{nt}^2} + \frac{2ab_k c_n (d_{kt} - a)}{d_{kk}^2 d_{nt}^2} \right) Z_{k-1} A^{-1} Z_{t-1} \]
\[
+ \frac{2}{T^2 (T-L)} \sum_{k,t=1}^{T} \left( \frac{b_k c_{kt}}{d_{kk}d_{nt}} - b_t \frac{a(c_{ks} - 1)^2 - 2b_k b_t (c_{ks} - 1) + b_k (b_k + b_t) c_{nt} + b_t^2 c_{kk}}{d_{kk}d_{nt}^2} \right) \]
\[
+ \frac{2}{T^2 (T-L)} \sum_{k,t=1}^{T} \left( \frac{b_k c_{nt}}{d_{kk}d_{nt}} - \frac{a b_k b_t c_{kt} + a^2 + a^2 c_{kk} c_{nt} - 2ad_{kt} + ac_{nt} d_{kt}}{d_{kk}^2 d_{nt}^2} \right) \]  
(A78)

Next is
\[ E[(\hat{\alpha}_1 - \alpha_1)(\hat{\alpha}_3 - \alpha_3)] \geq \frac{1}{T^2} \sum_{k,j=1}^{r} \left( \frac{2b_k^2}{(T-L)d_{kk}d_n} - a \frac{2a(c_{kk} - 1)^2 - 4b_k d_n}{(T-L)d_{kk}d_n} \frac{2b_k c_{kk}}{d_{kk}d_n} - \frac{2ad_{kk}}{(T-L)d_{kk}d_n} \right) \]

\[ + \frac{1}{T^2} \sum_{k,j=1}^{r} \left( a c_n \frac{-2ab_k b_{kl} + 2a^2 c_{kk} c_n - 2b_k^2 d_n - 2b_k d_{kk} - 4a d_{kl} + 2a c_{kl} d_{kl}}{(T-L)d_{kk}d_n^2} \right) \]

\[ + \frac{1}{T^2} \sum_{k,j=1}^{r} \left( \frac{4a(d_{kl} - a)}{T d_{kk}d_n^2} + \frac{ac_n}{T d_{kk}d_n^2} \right) Z_{k-1} A^{-1} Z_{t-1} \]

(A.79)

Continuing, we find

\[ E[(\hat{\alpha}_1 - \alpha_1)(\hat{\alpha}_3 - \alpha_3)] \geq \frac{1}{T^2} \sum_{k,j=1}^{r} \left( \frac{2b_k^2}{(T-L)d_{kk}d_n} - a \frac{2a(c_{kk} - 1)^2 - 4b_k d_n}{(T-L)d_{kk}d_n} \frac{2b_k c_{kk}}{d_{kk}d_n} - \frac{2ad_{kk}}{(T-L)d_{kk}d_n} \right) \]

\[ + \frac{1}{T^2} \sum_{k,j=1}^{r} \left( a c_n \frac{-2ab_k b_{kl} + 2a^2 c_{kk} c_n - 2b_k^2 d_n - 2b_k d_{kk} - 4a d_{kl} + 2a c_{kl} d_{kl}}{(T-L)d_{kk}d_n^2} \right) \]

\[ + \frac{4a}{T^2} \sum_{k,j=1}^{r} \left( \frac{a - d_{kl}}{T d_{kk}d_n^2} + \frac{ac_n(d_{kl} - a)}{T d_{kk}d_n^2} \right) Z_{k-1} A^{-1} Z_{t-1} \]
Selectively exchanging the indices, we find

\[
E[(\hat{\alpha}_1 - \alpha_1)(\hat{\alpha}_3 - \alpha_3)] \approx \frac{1}{T^2} \sum_{k,t=1}^{T} \left( \frac{2b^2_k - 2ac_{kk}}{(T-L)d_{kk}d_n} - a \frac{2a(c_{kk} - 1)^2 - 4b_kb_t(c_{kk} - 1) + 2b_k^2c_n}{(T-L)d_{kk}d_n^2} \right) + \frac{1}{T^2} \sum_{k,t=1}^{T} \left( \frac{-2ab_kb_t c_{kl} + 2a^2 + 2a^2 c_{kk}c_{tt} - 4a d_{kl} + 2a c_{kl}d_{kl}}{(T-L)d_{kk}d_n^2} \right) + \frac{4a}{T^2} \sum_{k,t=1}^{T} \left( \frac{(a - d_{kl})d_{kk}^2}{Td_{kk}^2d_n^2} + \frac{a c_{kk}(d_{kk}^2 - a c_{kk})}{Td_{kk}^2d_n^2} \right)Z'_{k-1}A^{-1}Z_{t-1} \]

\[
= \frac{1}{T^2} \sum_{k,t=1}^{T} \left( \frac{-2a d_{kk}}{(T-L)d_{kk}d_k} - a \frac{2a(c_{kk} - 1)^2 - 4b_kb_t(c_{kk} - 1) + 2b_k^2c_n}{(T-L)d_{kk}d_n^2} \right) + \frac{1}{T^2} \sum_{k,t=1}^{T} \left( \frac{-2a d_{kk}}{(T-L)d_{kk}d_k} - a \frac{2a(c_{kk} - 1)^2 - 4b_kb_t(c_{kk} - 1) + 2b_k^2c_n}{(T-L)d_{kk}d_n^2} \right) + \frac{4a}{T^2} \sum_{k,t=1}^{T} \left( \frac{(a - d_{kl})d_{kk}^2}{Td_{kk}^2d_n^2} \right)Z'_{k-1}A^{-1}Z_{t-1} \]

continuing, we find

\[
E[(\hat{\alpha}_1 - \alpha_1)(\hat{\alpha}_3 - \alpha_3)] \approx \frac{1}{T^2} \sum_{k,t=1}^{T} \left( \frac{-2a d_{kk}}{(T-L)d_{kk}d_k} - a \frac{2a(c_{kk} - 1)^2 - 4b_kb_t(c_{kk} - 1) + 2b_k^2c_n}{(T-L)d_{kk}d_n^2} \right) + \frac{2}{T^2} \sum_{k,t=1}^{T} \left( a^2 d_{kk} - b_k b_t c_{kl} + a + ac_{kk}c_{tt} - 2d_{kl} + c_{kl}d_{kl} \right) + \frac{4a}{T^2} \sum_{k,t=1}^{T} \left( \frac{(a - d_{kl})d_{kk}^2}{Td_{kk}^2d_n^2} \right)Z'_{k-1}A^{-1}Z_{t-1} \]

\[
(A.81)
\]

Recalling that \( \alpha_3 = \frac{a}{T} \sum_{t=1}^{T} \frac{1}{d_n} \), we have
Next is

\[
E\left[ (\hat{\alpha}_1 - \alpha_1) (\hat{\alpha}_3 - \alpha_3) \right] \equiv \frac{-2\alpha_3}{(T - L)a} - \frac{2a}{T^2(T - L)} \sum_{k,t=1}^r \frac{a(c_{kk} - 1)^2 - 2b_k b_t \left( c_{kk} - 1 \right) + b_t^2 c_{tt} + b_t^2 (c_{kk} + c_{tt})}{d_{kk}^2 d_{tt}^2} \\
+ \frac{2}{T^2(T - L)} \sum_{k,t=1}^r \left( a^2 c_{tt} - b_t b_t c_{tt} + a + ac_{kk} c_{tt} + (c_{tt} - 2) d_{tt} \right) - \frac{4a}{T^3} \sum_{k,t=1}^r \left( a - d_{tt} \right) b_t^2 \frac{Z_{t-1} A^{-1} Z_{t-1}}{d_{kk}^2 d_{tt}^2}
\]

(A.83)

where indices were exchanged in one of the last terms above. Continuing, we find

\[
E\left[ (\hat{\alpha}_2 - \alpha_2) (\hat{\alpha}_3 - \alpha_3) \right] \equiv \frac{1}{T^2} \sum_{k,t=1}^r \left( E\left[ (\hat{\alpha}_2 - \alpha_2) (\hat{\alpha}_3 - \alpha_3) \right] - \frac{E\left[ (\hat{a}_k - a_k) (\hat{d}_t - d_t) \right]}{d_{kk}^2 d_{tt}^2} \right) \\
+ \frac{1}{T^2} \sum_{k,t=1}^r \left( -b_j \frac{E\left[ (\hat{a}_k - a_k) (\hat{d}_t - d_t) \right]}{d_{kk}^2 d_{tt}^2} + ab_j \frac{E\left[ (\hat{d}_t - d_t) (\hat{d}_t - d_t) \right]}{d_{kk}^2 d_{tt}^2} \right)
\]

(A.84)
\[
E[(\hat{\alpha}_2 - \alpha_2)(\hat{\alpha}_3 - \alpha_3)] = \frac{2a^2}{T^2} \sum_{k,t=1}^{T} \left( -\frac{b_k b_t^2 c_{kt} + ab_t + ab_t c_{kt} c_{kt} - 2b_t d_{kt} + b_t c_{kt}}{(T - L) d_{kk}^2 a^2} \right) \\
+ \frac{1}{T^2} \sum_{k,t=1}^{T} \left( -\frac{2ab_t d_{kt}}{(T - L) d_{kk}^2 d_{tt}^2} + \frac{2ab_t^2 b_t d_{kt} - 2ab_t^3 d_{kk}^2}{(T - L) d_{kk}^2 d_{tt}^2} + \frac{4a^2 b_t (d_{kt} - a)}{T d_{kk}^2 d_{tt}^2} Z'_{k-1} A^{-1} Z_{t-1} \right) \\
= \frac{2a^2 b_t}{T^2} \sum_{k,t=1}^{T} \left( \frac{-b_k b_t c_{kt} + a + ac_{kt} c_{nt} - 2d_{kt} + c_{nt} d_{kt}}{(T - L) d_{kk}^2 d_n^2} \right) \\
+ \frac{1}{T^2} \sum_{k,t=1}^{T} \left( 2 \frac{-ab_t d_{nt} - ab_t^2 b_t - ab_t^3}{(T - L) d_{kk}^2 d_n^2} + \frac{4a^2 b_t (d_{nt} - a)}{T d_{kk}^2 d_n^2} Z'_{k-1} A^{-1} Z_{t-1} \right) \\
= \frac{2a^2 b_t}{T^2} \sum_{k,t=1}^{T} \left( \frac{a + ac_{kt} c_{nt} - b_k b_t c_{kt} + (c_{kt} - 2) d_{kt}}{(T - L) d_{kk}^2 d_n^2} \right) \\
+ \frac{1}{T^2} \sum_{k,t=1}^{T} \left( 2ab_t \frac{d_{nt} - b_t b_t = b_t^2}{(T - L) d_{kk}^2 d_n^2} + \frac{4a^2 b_t (d_{nt} - a)}{T d_{kk}^2 d_n^2} Z'_{k-1} A^{-1} Z_{t-1} \right) \\
= \frac{1}{T^2} \sum_{k,t=1}^{T} \left( 2a^2 \frac{b_t d_{nt}}{(T - L) d_{kk}^2 d_n^2} + 2a^2 b_t \frac{a + ac_{kt} c_{nt} - b_k b_t c_{kt} + (c_{kt} - 2) d_{kt}}{(T - L) d_{kk}^2 d_n^2} - 2ab_t \frac{b_t b_t + b_t^2}{(T - L) d_{kk}^2 d_n^2} \right) \\
+ \frac{1}{T^2} \sum_{k,t=1}^{T} \left( 4a^2 b_t (d_{nt} - a) \frac{Z'_{k-1} A^{-1} Z_{t-1}}{T d_{kk}^2 d_n^2} \right) \\
= \frac{1}{T^2} \sum_{k,t=1}^{T} \left( -2a \frac{b_t}{(T - L) d_{kk}^2 d_n^2} + 2a^2 b_t \frac{a + ac_{kt} c_{nt} - b_k b_t c_{kt} + (c_{kt} - 2) d_{kt}}{(T - L) d_{kk}^2 d_n^2} - 2ab_t \frac{b_t b_t + b_t^2}{(T - L) d_{kk}^2 d_n^2} \right) \\
+ \frac{1}{T^2} \sum_{k,t=1}^{T} \left( 4a^2 b_t (d_{nt} - a) \frac{Z'_{k-1} A^{-1} Z_{t-1}}{T d_{kk}^2 d_n^2} \right)
\]

We next make use of the facts that \( \Sigma b_t / d_{nt} = T \alpha_2 \) and \( \Sigma (1 / d_{nt}) = \alpha_3 T / a \) so that
\[
\frac{1}{T^2} \sum_{k,t=1}^{T} \left( -2a \frac{b_t}{(T-L)d_{kk}d_{tt}} \right) = \frac{-2a}{T^2(T-L)} \left( \sum_{k=1}^{T} \frac{1}{d_{kk}} \right) \left( \sum_{t=1}^{T} \frac{b_t}{d_{tt}} \right) = \frac{-2a}{T^2(T-L)} \left( \alpha_3 T \right) (T \alpha_2) = \frac{-2\alpha_3 \alpha_2}{T-L}
\]

which we use to find

\[
E \left[ (\hat{\alpha}_2 - \alpha_2)(\hat{\alpha}_3 - \alpha_3) \right] \approx \frac{-2\alpha_3 \alpha_2}{(T-L)} + \frac{4a^2}{T^3} \sum_{k,t=1}^{T} \frac{b_t(d_{kk} - a)}{d_{kk}^2 d_{tt}^2} Z_{k-1}^T A^{-1} Z_{t-1}
\]

\[
+ \frac{2a}{T^2(T-L)} \sum_{k,t=1}^{T} \left( ab_t \frac{a + ac_{kk}c_{tt} - b_k b_t c_{kk}}{d_{kk}^2 d_{tt}^2} - \frac{b_t^2(b_k + b_t)}{d_{kk}^2 d_{tt}^2} \right)
\]

completing the proof. \(\square\)

**Lemmas for Proposition 11**

**Lemma 1:** Let \(u\) be a vector. Then

\[
(I + uu')^{-1} = I - \frac{uu'}{1 + uu'}
\]

**Proof:** We multiply:

\[
(I + uu') \left( I - \frac{uu'}{1 + uu'} \right) = \left( I - \frac{uu'}{1 + uu'} \right) + uu' \left( I - \frac{uu'}{1 + uu'} \right)
\]

\[
= I - \frac{uu'}{1 + uu'} + uu' - \frac{u(uu')u'}{1 + uu'} = I + uu' \left( -\frac{1}{1 + uu'} + 1 - \frac{uu'}{1 + uu'} \right) = I
\]

completing the proof. \(\square\)

**Lemma 2:** Let \(u\) be a vector and let \(V\) be a fixed positive semidefinite matrix. Then

\[
(V + uu')^{-1} = V^{-1} - \frac{V^{-1}uu'V^{-1}}{1 + uu'V^{-1}u}
\]

**Proof:** Let \(V = AA'\) be the Cholesky decomposition. Then we have:
\[(V + uu')^{-1} = (AA' + uu')^{-1} = (A')^{-1} \left[ I + (A^{-1}u)(A^{-1}u)' \right]^{-1} A^{-1} \]  \hspace{1cm} (A.91)

Applying Lemma 1, we find

\[
(V + uu')^{-1} = (A')^{-1} \left\{ I - \frac{(A^{-1}u)(A^{-1}u)'}{1 + (A^{-1}u)'(A^{-1}u)} \right\} A^{-1}
\]

\[
= (A')^{-1} A^{-1} - \frac{(A^{-1}u)(A^{-1}u)'}{1 + (A^{-1}u)'(A^{-1}u)} A^{-1}
\]

\[
= (AA')^{-1} - \frac{(AA')^{-1} uu'(AA')^{-1}}{1 + u'(AA')^{-1} u} = \lambda^{-1} - \frac{V^{-1} uu' V^{-1}}{1 + u' V^{-1} u}
\]  \hspace{1cm} (A.92)

completing the proof. □

**Lemma 3:** Here are some expressions with \( \Lambda(Z) = \left[ V(Z) + \mu(Z) \mu'(Z) \right]^{-1} \). For simplicity we will omit the argument \( Z \).

\[
\Lambda = V^{-1} - V^{-1} \mu' V^{-1} \frac{1}{1 + \mu' V^{-1} \mu}, \quad \Lambda \mu = \frac{V^{-1} \mu}{1 + \mu' V^{-1} \mu}, \quad \text{and} \quad \mu' \Lambda \mu = 1 - \frac{1}{1 + \mu' V^{-1} \mu}
\]  \hspace{1cm} (A.93)

**Proof:** Using Lemma 2, we find

\[
\Lambda = V^{-1} - V^{-1} \mu' V^{-1} \frac{1}{1 + \mu' V^{-1} \mu}
\]  \hspace{1cm} (A.94)

Next, we compute:

\[
\Lambda \mu = \left( V^{-1} - V^{-1} \mu' V^{-1} \right) \mu = V^{-1} \mu - V^{-1} \mu' V^{-1} \mu = \left( 1 - \frac{\mu' V^{-1} \mu}{1 + \mu' V^{-1} \mu} \right) V^{-1} \mu
\]

\[
= \left( \frac{1 + \mu' V^{-1} \mu - \mu' V^{-1} \mu}{1 + \mu' V^{-1} \mu} \right) V^{-1} \mu = \frac{V^{-1} \mu}{1 + \mu' V^{-1} \mu}
\]  \hspace{1cm} (A.95)
using this we find
\[ \mu' \Lambda \mu = \mu' \frac{V^{-1} \mu}{1 + \mu' V^{-1} \mu} = \frac{1 + \mu' V^{-1} \mu - 1}{1 + \mu' V^{-1} \mu} = 1 - \frac{1}{1 + \mu' V^{-1} \mu} \]
(A.96)

completing the proof. \( \square \)

**Lemma 4:** Let \( W \) be a small perturbation of \( V \). Then
\[ (V + W)^{-1} = V^{-1} - V^{-1} W V^{-1} + V^{-1} W W^{-1} V^{-1} + O(\|W\|^3) \]  
(A.97)

and
\[ \hat{V}^{-1} \cong V^{-1} - V^{-1} (\hat{V} - V) V^{-1} + V^{-1} (\hat{V} - V) V^{-1} (\hat{V} - V) V^{-1} \]  
(A.98)

**Proof:** we multiply
\[
(V + W)(V^{-1} - V^{-1} W V^{-1} + V^{-1} W W^{-1} V^{-1}) \\
= V(V^{-1} - V^{-1} W V^{-1} + V^{-1} W W^{-1} V^{-1}) + W(V^{-1} - V^{-1} W V^{-1} + V^{-1} W W^{-1} V^{-1}) \\
= (I - W V^{-1} + W V^{-1} W V^{-1}) + (W V^{-1} - W V^{-1} W V^{-1} + W V^{-1} W V^{-1} W V^{-1}) \\
= I + W V^{-1} W V^{-1} - W V^{-1} + W V^{-1} W V^{-1} - W V^{-1} W V^{-1} + W V^{-1} W V^{-1} W V^{-1} \\
= I + W V^{-1} W V^{-1} W V^{-1} = I + O(\|W\|^3) 
\]  
(A.99)

Note that we may apply this result to obtain
\[ \hat{V}^{-1} = \left[ V + (\hat{V} - V) \right]^{-1} \cong V^{-1} - V^{-1} (\hat{V} - V) V^{-1} + V^{-1} (\hat{V} - V) V^{-1} (\hat{V} - V) V^{-1} \]  
(A.100)

completing the proof. \( \square \)

**Lemma 5:** \( E[(\hat{\mu}_k - \mu_k)' Q (\hat{\mu}_i - \mu_i)] = \frac{tr(VQ)}{T} Z_{k,i} A^{-1} Z_{i,-1} \) and, in particular,
\[
E\left( (\hat{\mu}_k - \mu_k)' V^{-1} (\hat{\mu}_t - \mu_t) \right) = \frac{n}{T} Z'_{k-1} A^{-1} Z_{t-1}, \quad E\left( (\hat{\mu}_k - \mu_k)(\hat{\mu}_t - \mu_t)' \right) = \frac{V Z'_{k-1} A^{-1} Z_{t-1}}{T} \quad (A.101)
\]

and
\[
E\left( (\hat{\mu}_k - \mu_k)' V^{-1} u' V^{-1} (\hat{\mu}_t - \mu_t) \right) = \frac{u' V^{-1} u}{T} Z'_{t-1} A^{-1} Z_{t-1} \quad (A.102)
\]

Proof: Recall that \( \hat{\delta} - \delta = A^{-1} \left( \frac{1}{T} \sum_{t=1}^{T} Z_{t-1} \varepsilon_t' \right) \), so that
\[
\hat{\mu}_t - \mu_t = (\hat{\delta} - \delta)' Z_{t-1} = \left( \frac{1}{T} \sum_{k=1}^{T} \varepsilon_k Z'_{k-1} \right) A^{-1} Z_{t-1} \quad (A.103)
\]

Using commutativity within the trace operator, independence of \( (\varepsilon_t, \varepsilon_k) \) when \( k \neq t \), and taking advantage of commutativity of scalars, we find:
\[
E\left( (\hat{\mu}_k - \mu_k)' Q(\hat{\mu}_t - \mu_t) \right) = E\left( \text{tr} \left( (\hat{\mu}_k - \mu_k)' Q(\hat{\mu}_t - \mu_t) \right) \right) = E\left( \text{tr} \left( (\hat{\mu}_t - \mu_t) (\hat{\mu}_k - \mu_k)' Q \right) \right)
\]
\[
= \text{tr} \left( E\left( (\hat{\mu}_t - \mu_t) (\hat{\mu}_k - \mu_k)' Q \right) \right) = \text{tr} \left( E\left( \left( \frac{1}{T} \sum_{t=1}^{T} \varepsilon_t Z'_{t-1} \right) A^{-1} Z_{t-1} Z'_{k-1} A^{-1} \left( \frac{1}{T} \sum_{t=1}^{T} \varepsilon_t Z'_{t-1} \right)^T \right) \right)
\]
\[
= \text{tr} \left( E\left( \frac{1}{T^2} \sum_{t=1}^{T} \varepsilon_t Z'_{t-1} A^{-1} Z_{t-1} Z'_{k-1} A^{-1} Z_{t-1} \varepsilon_t' \right) \right)
\]
\[
= \frac{1}{T^2} \sum_{t=1}^{T} \text{tr} \left( E(\varepsilon_t \varepsilon_t') Z'_{t-1} A^{-1} Z_{t-1} Z'_{k-1} A^{-1} Z_{t-1} \right)
\]
\[
= \frac{1}{T^2} \sum_{t=1}^{T} \text{tr} \left( (VQ) Z'_{t-1} A^{-1} Z_{t-1} Z'_{k-1} A^{-1} Z_{t-1} \right) = \frac{\text{tr}(VQ)}{T^2} \sum_{t=1}^{T} Z'_{t-1} A^{-1} Z_{t-1} Z'_{k-1} A^{-1} Z_{t-1}
\]
\[
= \frac{\text{tr}(VQ)}{T^2} Z'_{k-1} A^{-1} \left( \sum_{t=1}^{T} Z_{t-1} Z'_{t-1} \right) A^{-1} Z_{t-1} = \frac{\text{tr}(VQ)}{T} Z'_{k-1} A^{-1} A A^{-1} Z_{t-1}
\]
\[
= \frac{\text{tr}(VQ)}{T} Z'_{k-1} A^{-1} Z_{t-1}
\]
where we note that $Z_{t-1}'A^{-1}Z_{t-1} = Z_{t-1}'A^{-1}Z_{t-1}$ because each is a transposed scalar of the other. In particular, we have

$$E\left[\left(\hat{\mu}_k - \mu_k\right)'V^{-1}\left(\hat{\mu}_k - \mu_k\right)\right] = \frac{tr(VV^{-1})}{T}Z_{k-1}'A^{-1}Z_{t-1} = \frac{tr(I_n)}{T}Z_{k-1}'A^{-1}Z_{t-1} = \frac{n}{T}Z_{k-1}'A^{-1}Z_{t-1}$$  \hspace{1cm} (A.105)

Moreover, for any fixed vectors $x$ and $y$, we have

$$x'E\left[\left(\hat{\mu}_k - \mu_k\right)(\hat{\mu}_k - \mu_k)'\right]y = E\left[x'(\hat{\mu}_k - \mu_k)(\hat{\mu}_k - \mu_k)'y\right] = E\left[\left(\hat{\mu}_k - \mu_k\right)'xy'(\hat{\mu}_k - \mu_k)\right] = \frac{tr(Vxy')}{T}Z_{k-1}'A^{-1}Z_{t-1} = \frac{x'V'y}{T}Z_{k-1}'A^{-1}Z_{t-1} = \frac{x'VZ_{k-1}'A^{-1}Z_{t-1}}{T}y$$  \hspace{1cm} (A.106)

which states equality for all $x$ and $y$, implying that

$$E\left[\left(\hat{\mu}_k - \mu_k\right)(\hat{\mu}_k - \mu_k)'\right] = \frac{VZ_{k-1}'A^{-1}Z_{t-1}}{T}$$  \hspace{1cm} (A.107)

and

$$E\left[\left(\hat{\mu}_k - \mu_k\right)'V^{-1}u_2'V^{-1}(\hat{\mu}_k - \mu_k)\right] = \frac{tr(VV^{-1}u_2'V^{-1})}{T}Z_{k-1}'A^{-1}Z_{t-1} = \frac{u_2'V^{-1}u_2}{T}Z_{k-1}'A^{-1}Z_{t-1}$$  \hspace{1cm} (A.108)

and completing the proof. □

**Lemma 6**: For a fixed matrix $Q$, we have

$$E\left[(\hat{V} - V)Q(\hat{V} - V)\right] = \frac{VQ'V + V\left[tr(VQ)\right]}{T - L}$$  \hspace{1cm} (A.109)

and, in particular,
\[ E\left[(\hat{V} - V)V^{-1}(\hat{V} - V)\right] = \left(\frac{n+1}{T-L}\right) \]  
\[ \text{as well as} \]
\[ E\left[u'_i V^{-1}(\hat{V} - V)V^{-1}u_i' V^{-1}(\hat{V} - V)V^{-1}u_i\right] = \frac{u'_i V^{-1}u_i' V^{-1}u_i + u'_i V^{-1}u_i' V^{-1}u_i}{T-L} \]

which implies that:

\[ \mu'_i V^{-1}E\left[(\hat{V} - V)V^{-1}\mu_i \mu'_i V^{-1}(\hat{V} - V)\right]V^{-1}\mu_i = \frac{2(c_u - 1)^2}{T-L} \]

\[ 1'V^{-1}E\left[(\hat{V} - V)V^{-1}11V^{-1}(\hat{V} - V)\right]V^{-1}1 = \frac{2a^2}{T-L} \]

\[ \mu'_i V^{-1}E\left[(\hat{V} - V)V^{-1}11V^{-1}(\hat{V} - V)\right]V^{-1}1 = \frac{2ab_i}{T-L} \]

\[ \mu'_i V^{-1}E\left[(\hat{V} - V)V^{-1}\mu_i \mu'_i V^{-1}(\hat{V} - V)\right]V^{-1}1 = \frac{2b_i (c_u - 1)}{T-L} \]  
\[ 1'V^{-1}E\left[(\hat{V} - V)V^{-1}1\mu_i V^{-1}(\hat{V} - V)\right]V^{-1}1 = \frac{a (c_u - 1) + b_i b_i}{T-L} \]

\[ \mu'_i V^{-1}E\left[(\hat{V} - V)V^{-1}\mu_i 1' V^{-1}(\hat{V} - V)\right]V^{-1}1 = \frac{2b_i^2}{T-L} \]

\[ \mu'_i V^{-1}E\left[(\hat{V} - V)V^{-1}\mu_i \mu'_i V^{-1}(\hat{V} - V)\right]V^{-1}\mu_i = \frac{2(c_k - 1)^2}{T-L} \]

where many others of the form \( E[u'_i V^{-1}(\hat{V} - V)V^{-1}u_i' V^{-1}(\hat{V} - V)V^{-1}u_i] \) may be transformed (commutativity of scalar multiplication, transposing a scalar) into one from the above list. Those having four repeated instances with each \( u_i \in \{\mu_i, 1\} \) are covered in the first two in the above list. Those having exactly three repeated instances with each \( u_i \in \{\mu_i, 1\} \) along with one of the other type are covered in the next two in the list. Those having two of one type and two of the other with each \( u_i \in \{\mu_k, \mu_i, 1\} \) are covered by the fifth and sixth in the list.
Proof: While Haff (1979) Theorem 3.1(iii) can be used to prove this in the case of a positive semidefinite $Q$, we need the more general result for cases such as $Q = V^{-1} \mu_i 1' V^{-1} = (V^{-1} \mu_i) (V^{-1} 1)'$ with $\mu_i$ and $1$ noncollinear, for which $Q$ is not positive semidefinite because when $Q = xy'$ with noncollinear $x$ and $y$, we can find a vector $u$ such that $u'Qu = u'xy'u < 0$, for example by choosing $u = x + a \left( y - \frac{xy'}{xx'} \right)$ with scalar $a$, for which

$$u'Qu = u'xy'u = \left[ x + a \left( y - \frac{xy'}{xx'} \right) \right]' y y' \left[ x + a \left( y - \frac{xy'}{xx'} \right) \right]$$

$$= x'xy'x + ax' \left( y'y - \frac{xy'}{xx'} y'x \right) + a \left( y'x - \frac{xy'}{xx'} x'x \right) y'x + a^2 \left( y - \frac{xy'}{xx'} \right)' y y' \left( y - \frac{xy'}{xx'} \right) \quad \text{(A.113)}$$

in which the coefficient of $a$ is positive and cannot be zero because $x$ and $y$ are assumed to be noncollinear (by the Cauchy-Schwartz Inequality). Choosing a negative $a$ with large magnitude shows that such a $Q$ is not positive semidefinite.

We next prove a related result when $V = I$:

$$E(YQY) = p(p + 1)Q + p \left[ tr(Q) \right] I - p(Q - Q') \quad \text{(A.114)}$$

for an $n \times n$ matrix $Y$ with the Wishart distribution $W(n, p, I)$, and $Q$ an arbitrary $n \times n$ matrix of constants. We may represent $Y$ using $X_{it} \overset{iid}{\sim} N(0,1)$ for $i = 1...n$ and $t = 1,...,p$, defining $Y_{ij} = \sum_{t=1}^{p} X_{it} X_{jt}$, for which the matrix $Y$ has a Wishart distribution $W(n, p, I)$. Note that

$$\left[ E(YQY) \right]_{ij} = E\left( \sum_{k,l=1}^{n} Y_{ik} Q_{kl} Y_{lj} \right) = \sum_{k,l=1}^{n} Q_{kl} E\left( Y_{ik} Y_{lj} \right) \quad \text{(A.115)}$$
so we need \( E(Y_{iu}Y_{iu}) \). There are four cases:

- All four subscripts identical so that \( i = j = k = l \): In this case, \( Y_{ii} = \sum_{t=1}^{p} X_{it}^2 \) and

\[
E(Y_{ii}^2) = E\left[\left(\sum_{t=1}^{p} X_{it}^2\right)^2\right] = \sum_{t,u=1}^{p} E\left(X_{it}^2 X_{iu}^2\right)
\]

(A.116)

When \( t \neq u \), \( X_{it}^2 \) and \( X_{iu}^2 \) are independent and we have

\[
E(Y_{ii}^2) = \sum_{t,u=1}^{p} E\left(X_{it}^2 X_{iu}^2\right) = pE(X_{ii}^4) + p(p-1)E(X_{ii}^2 X_{2i}) = 3p + p(p-1) = p(p + 2)
\]

(A.117)

- Exactly three identical, for example, \( i = j = k \) with \( l \neq i \)

\[
E(Y_{ik}Y_{lk}) = E(Y_{ik}Y_{ki}) = E\left[\left(\sum_{t=1}^{p} X_{it}X_{lt}\right)^2\right] = \sum_{t,u=1}^{p} E\left(X_{it}X_{lt}X_{iu}X_{lu}\right) = 0
\]

because \( X_{iu} \) is independent of both \( X_{it} \) and \( X_{iu} \).

- Exactly two identical pairs of the form \( E(Y_{ik}Y_{ki}) = E\left(Y_{ik}^2\right) \) with \( i \neq k \). We find

\[
E(Y_{ik}Y_{ki}) = E\left[\left(\sum_{t=1}^{p} X_{it}X_{kt}\right)^2\right] = \sum_{t,u=1}^{p} E\left(X_{it}X_{kt}X_{iu}X_{ku}\right)
\]

(A.119)

When \( t \neq u \) we have \( \sum_{t,u=1}^{p} E(X_{it}X_{kt})E(X_{iu}X_{ku}) = 0 \) by independence.

We therefore have

\[
E(Y_{ik}Y_{ki}) = E\left(Y_{ik}^2\right) = \sum_{t=1}^{p} E\left(X_{it}X_{kt}X_{iu}X_{ku}\right) = \sum_{t=1}^{p} E\left(X_{it}^2 X_{kt}^2\right) = p
\]

(A.120)

- Exactly two identical pairs of the form \( E(Y_{ii}Y_{jj}) \) with \( i \neq j \), for which

\[
E(Y_{ii}Y_{jj}) = E\left[\left(\sum_{t=1}^{p} X_{it}^2\right)\left(\sum_{t=1}^{p} X_{jt}^2\right)\right] = E\left[\sum_{t=1}^{p} X_{it}^2\right]\left[\sum_{t=1}^{p} X_{jt}^2\right] = p^2
\]

(A.121)
• All other cases have expectation zero by independence of any single, non-repeated, subscript term.

There are two cases for evaluating $E(YQY)_{ij} = \sum_{k,l=1}^{n} Q_{kl} E(Y_k Y_l)$

- If $i = j$ then
  \[
  E(YQY)_{ii} = \sum_{k,l=1}^{n} Q_{kl} E(Y_k Y_l) = \sum_{k=1}^{n} Q_{kk} E(Y_k Y_k) + \sum_{k \neq l} Q_{kl} E(Y_k Y_l)
  \]
  \[
  = \sum_{k \neq l} Q_{kk} E(Y_k Y_k) + Q_{ii} E(Y_i^2) + \sum_{k \neq l} Q_{ll} E(Y_l Y_l)
  \]
  \[
  = p \sum_{k \neq l} Q_{kk} + p(p + 2)Q_{ii} + 0 = p \sum_{k \neq l} Q_{kk} + pQ_{ii} + p(p + 1)Q_{ii}
  \]
  \[
  = p[tr(Q)] + p(p + 1)Q_{ii}
  \]  

- If $i \neq j$ then the only nonzero expectations $E(Y_k Y_l)$ occur either when $k = i$ and $l = j$, or when $k = j$ and $l = i$. We find
  \[
  E(YQY)_{ij} = \sum_{k,l=1}^{n} Q_{kl} E(Y_k Y_l) = Q_{ii} E(Y_i Y_j) + Q_{jj} E(Y_j^2)
  \]
  \[
  = Q_{ij} E(Y_i Y_j) + Q_{ji} E(Y_j^2) = p^2 Q_{ij} + pQ_{ji}
  \]  

and it will be helpful to have an alternative representation:

\[
E(YQY)_{ij} = p^2 Q_{ij} + pQ_{ji} = p(p + 1)Q_{ij} - pQ_{ij} + pQ_{ji} = p(p + 1)Q_{ij} - p(Q_{ij} - Q_{ji})
\]  

(A.124)

Putting these together, we find

\[
E(YQY) = p(p + 1)Q + p[tr(Q)]I - p(Q - Q')
\]  

(A.125)

for $Y \sim W(n, p, I)$ with general $Q$.

Now we consider general $V$ and note that $(T - L)V \sim W(n, T - L, V)$. Choose $n \times n$ matrix $A$ using the Cholesky Decomposition such that $AVA' = I$, $V = A^{-1}A'^{-1}$ and $V^{-1} = A'A$. Then
\((T - L)A\hat{V}A' \sim W(n, T - L, A\hat{V}A') = W(n, T - L, I)\) and \((T - L)A\hat{V}A^d = Y\) with \(p = T - L\) and \(Y \sim W(n, p, I)\). We then have

\[
E(\hat{V}QV) = \frac{1}{(T - L)^2} E\left\{A^{-1}\left[(T - L)A\hat{V}A'(A^{-1}QA^{-1})(T - L)A\hat{V}A'\right]A^{-1}\right\}
\]

\[
= \frac{1}{(T - L)^2} E\left\{A^{-1}\left[Y\left(A^{-1}QA^{-1}\right)Y\right]A^{-1}\right\}
\]

\[
= \frac{1}{(T - L)^3} A^{-1}\left\{(T - L)(T - L + 1)A^{-1}QA^{-1} + (T - L)\left[tr\left(A^{-1}QA^{-1}\right)\right]I - (T - L)\left[A^{-1}QA^{-1} - (A^{-1}QA^{-1})'\right]\right\}A^{-1}
\]

\[
= \frac{1}{(T - L)}\left[(T - L + 1)A^{-1}QA^{-1}A^{-1} + tr\left(QA^{-1}A^{-1}\right)A^{-1}A^{-1} - A^{-1}QA^{-1}A^{-1} - A^{-1}QA^{-1}QA^{-1}A^{-1}\right]
\]

\[
= \frac{1}{(T - L)}\left[(T - L + 1)VQV + tr\left(QV\right)V - V\left(Q - Q'\right)V\right]
\]

(A.126)

and therefore

\[
E\left[\left(\hat{V} - V\right)Q\left(\hat{V} - V\right)\right] = E(\hat{V}QV) - VQV
\]

\[
= \frac{1}{(T - L)}\left[(T - L + 1)VQV + tr\left(QV\right)V - V\left(Q - Q'\right)V\right] - VQV
\]

\[
= VQV + \frac{1}{(T - L)}VQV + \frac{1}{(T - L)}tr\left(QV\right)V - \frac{1}{(T - L)}V\left(Q - Q'\right)V - VQV
\]

\[
= \frac{VQV + V\left[tr\left(QV\right)\right] - V\left(Q - Q'\right)V}{(T - L)} = \frac{VQ'V + V\left[tr\left(QV\right)\right]}{(T - L)}
\]

(A.127)

completing the proof of the first result. The special cases then follow immediately from this general result. For example,
\[
E[u'V^{-1}(\hat{V} - V)V^{-1}u'Y^{-1}V^{-1}u] = u'V^{-1}E[(\hat{V} - V)V^{-1}u'V^{-1}V^{-1}u]
\]
\[
= u'V^{-1}V^{-1}u'V^{-1}V^{-1}u + V[tr(VV^{-1}u'Y^{-1}u)]V^{-1}u
\]
\[
= \frac{u'V^{-1}V^{-1}u'V^{-1}V^{-1}u + u'V^{-1}u[tr(u'Y^{-1}u)\right]}{T - L}
\]
\[
= \frac{u'V^{-1}u'V^{-1}u + u'V^{-1}u'Y^{-1}u}{T - L}
\]
 Alonso, completing the proof. □

**Lemma 7:** Rewriting the portfolio constants using the definitions above we find their consistent estimators as:

\[
\hat{\alpha}_1 = \frac{1}{T} \sum_{t=1}^{T} \frac{1 + \hat{\mu}'Y^{-1}\hat{\mu}}{1\hat{Y}'1(1 + \hat{\mu}'Y^{-1}\hat{\mu}) - (1\hat{Y}'\hat{\mu})^2} = \frac{1}{T} \sum_{t=1}^{T} \hat{c}_t
\]

\[
\hat{\alpha}_2 = \frac{1}{T} \sum_{t=1}^{T} \frac{1\hat{Y}'\hat{\mu}}{1\hat{Y}'1(1 + \hat{\mu}'Y^{-1}\hat{\mu}) - (1\hat{Y}'\hat{\mu})^2} = \frac{1}{T} \sum_{t=1}^{T} \hat{b}_t
\]

\[
\hat{\alpha}_3 = \frac{1}{T} \sum_{t=1}^{T} \frac{1\hat{Y}'1}{1\hat{Y}'1(1 + \hat{\mu}'Y^{-1}\hat{\mu}) - (1\hat{Y}'\hat{\mu})^2} = \frac{1}{T} \sum_{t=1}^{T} \frac{1}{\hat{d}_t}
\]

Proof: Using Lemma 2, we find

\[
\Lambda = V^{-1} - V^{-1}\frac{\mu'Y^{-1}}{1 + \mu'Y^{-1}\mu} = \frac{1}{T} \sum_{t=1}^{T} \frac{1}{\hat{d}_t}
\]

from which we find expressions for the portfolio constants in terms of \( \mu \) and \( V \).
\[ \alpha_1 = E \left( \frac{1}{1'\Lambda_1} \right) = E \left( \frac{1}{1' \left( V^{-1} - \frac{V^{-1} \mu V^{-1}}{1 + \mu' V^{-1} \mu} \right)} \right) = E \left( \frac{1}{1' V^{-1} - \frac{V^{-1} \mu V^{-1}}{1 + \mu' V^{-1} \mu}} \right) \]

\[ = E \left( \frac{1 + \mu' V^{-1} \mu}{1' V^{-1} \left( 1 + \mu' V^{-1} \mu \right) - \left( 1' V^{-1} \mu \right)^2} \right) \]

\[ \alpha_2 = E \left( \frac{1' \Lambda_\mu}{1' \Lambda_1} \right) = E \left( \frac{1' V^{-1} \mu - \frac{1' V^{-1} \mu V^{-1} \mu}{1 + \mu' V^{-1} \mu}}{1' V^{-1} - \frac{1' V^{-1} \mu V^{-1} \mu}{1 + \mu' V^{-1} \mu}} \right) = E \left( \frac{1' V^{-1} \mu \left( 1 + \mu' V^{-1} \mu \right) - 1' V^{-1} \mu V^{-1} \mu \mu'}{1' V^{-1} \left( 1 + \mu' V^{-1} \mu \right) - \left( 1' V^{-1} \mu \right)^2} \right) \]

\[ = E \left( \frac{1' V^{-1} \mu \left( 1 + \mu' V^{-1} \mu \right) - 1' V^{-1} \mu V^{-1} \mu \mu'}{1' V^{-1} \left( 1 + \mu' V^{-1} \mu \right) - \left( 1' V^{-1} \mu \right)^2} \right) \]
\[ \alpha_3 = 1 - E\left[ \mu' \left( \Lambda - \frac{\Lambda \Xi' \Lambda}{\Xi' \Xi} \right) \mu \right] = 1 - E \left( \mu' \Lambda \mu - \frac{\mu' \Lambda \Xi' \Lambda \mu}{\Xi' \Xi} \right) \]
\[ = 1 - E \left( \mu' V^{-1} \mu - \frac{\mu' V^{-1} \mu' V^{-1} \mu}{1 + \mu' V^{-1} \mu} \right) \frac{1}{1 + \mu' V^{-1} \mu} \left( \frac{1}{1 + \mu' V^{-1} \mu} \right)^2 \]
\[ = 1 - E \left( \frac{\mu' V^{-1} \mu}{1 + \mu' V^{-1} \mu} \right) \frac{1}{1 + \mu' V^{-1} \mu} \left( \frac{1}{1 + \mu' V^{-1} \mu} \right)^2 \]
\[ = E \left( \frac{1}{1 + \mu' V^{-1} \mu} \right) \frac{1}{1 + \mu' V^{-1} \mu} \left( \frac{1}{1 + \mu' V^{-1} \mu} \right)^2 \]
\[ = E \left( \frac{1}{1 + \mu' V^{-1} \mu} \right) \frac{1}{1 + \mu' V^{-1} \mu} \left( \frac{1}{1 + \mu' V^{-1} \mu} \right)^2 \]
\[ = E \left( \frac{1}{1 + \mu' V^{-1} \mu} \right) \frac{1}{1 + \mu' V^{-1} \mu} \left( \frac{1}{1 + \mu' V^{-1} \mu} \right)^2 \]
\[ = E \left( \frac{1}{1 + \mu' V^{-1} \mu} \right) \frac{1}{1 + \mu' V^{-1} \mu} \left( \frac{1}{1 + \mu' V^{-1} \mu} \right)^2 \]
\[ = E \left( \frac{1}{1 + \mu' V^{-1} \mu} \right) \frac{1}{1 + \mu' V^{-1} \mu} \left( \frac{1}{1 + \mu' V^{-1} \mu} \right)^2 \]
\[ = E \left( \frac{1}{1 + \mu' V^{-1} \mu} \right) \frac{1}{1 + \mu' V^{-1} \mu} \left( \frac{1}{1 + \mu' V^{-1} \mu} \right)^2 \]
\[ = E \left( \frac{1}{1 + \mu' V^{-1} \mu} \right) \frac{1}{1 + \mu' V^{-1} \mu} \left( \frac{1}{1 + \mu' V^{-1} \mu} \right)^2 \]
\[ = E \left( \frac{1}{1 + \mu' V^{-1} \mu} \right) \frac{1}{1 + \mu' V^{-1} \mu} \left( \frac{1}{1 + \mu' V^{-1} \mu} \right)^2 \]
\[ = E \left( \frac{1}{1 + \mu' V^{-1} \mu} \right) \frac{1}{1 + \mu' V^{-1} \mu} \left( \frac{1}{1 + \mu' V^{-1} \mu} \right)^2 \]
\[ = E \left( \frac{1}{1 + \mu' V^{-1} \mu} \right) \frac{1}{1 + \mu' V^{-1} \mu} \left( \frac{1}{1 + \mu' V^{-1} \mu} \right)^2 \]
\[ = E \left( \frac{1}{1 + \mu' V^{-1} \mu} \right) \frac{1}{1 + \mu' V^{-1} \mu} \left( \frac{1}{1 + \mu' V^{-1} \mu} \right)^2 \]
\[ = E \left( \frac{1}{1 + \mu' V^{-1} \mu} \right) \frac{1}{1 + \mu' V^{-1} \mu} \left( \frac{1}{1 + \mu' V^{-1} \mu} \right)^2 \]
\[ = E \left( \frac{1}{1 + \mu' V^{-1} \mu} \right) \frac{1}{1 + \mu' V^{-1} \mu} \left( \frac{1}{1 + \mu' V^{-1} \mu} \right)^2 \]
\[ = E \left( \frac{1}{1 + \mu' V^{-1} \mu} \right) \frac{1}{1 + \mu' V^{-1} \mu} \left( \frac{1}{1 + \mu' V^{-1} \mu} \right)^2 \]
\[ = E \left( \frac{1}{1 + \mu' V^{-1} \mu} \right) \frac{1}{1 + \mu' V^{-1} \mu} \left( \frac{1}{1 + \mu' V^{-1} \mu} \right)^2 \]
\[ = E \left( \frac{1}{1 + \mu' V^{-1} \mu} \right) \frac{1}{1 + \mu' V^{-1} \mu} \left( \frac{1}{1 + \mu' V^{-1} \mu} \right)^2 \]
\[ = E \left( \frac{1}{1 + \mu' V^{-1} \mu} \right) \frac{1}{1 + \mu' V^{-1} \mu} \left( \frac{1}{1 + \mu' V^{-1} \mu} \right)^2 \]

The estimates make use of the estimated values \( \hat{V} \) and \( \hat{\mu} \). \( \square \)

**Lemma 8**: Second-order expansions may be found as follows, omitting terms involving both \( \hat{\mu} - \mu \) and \( \hat{V} - V \) because they are independent of one another and will have expectation zero. Keeping only relevant terms up to second order, we find:

\[ \hat{\alpha} - \alpha \approx -1V^{-1}(\hat{V} - V)V^{-1}1 + 1V^{-1}(\hat{V} - V)V^{-1}(\hat{V} - V)V^{-1}1 \]
\[ (\hat{\alpha} - \alpha)^2 \approx 1V^{-1}(\hat{V} - V)V^{-1}1V^{-1}(\hat{V} - V)V^{-1}1 \]
\[
\hat{b}_i - b_i \equiv -1V^{-1}(\hat{V} - V)V^{-1}\mu_i + 1V^{-1}(\hat{\mu}_r - \mu_r) + 1V^{-1}(\hat{V} - V)V^{-1}(\hat{V} - V)V^{-1}\mu_i
\]

\[
(\hat{b}_k - b_k)(\hat{b}_k - b_k) \equiv 1V^{-1}(\hat{V} - V)V^{-1}\mu_k 1V^{-1}(\hat{V} - V)V^{-1}\mu_i + (\hat{\mu}_k - \mu_k)^T V^{-1} 11V^{-1}(\hat{\mu}_r - \mu_r)
\]

\[
(\hat{a} - a)(\hat{b}_i - b_i) \equiv 1V^{-1}(\hat{V} - V)V^{-1} 11V^{-1}(\hat{V} - V)V^{-1}\mu_i
\]

\[
\hat{c}_n - c_n \equiv -\mu_i' V^{-1}(\hat{V} - V)V^{-1}\mu_i + 2\mu_i' V^{-1}(\hat{\mu}_r - \mu_r)
\]

\[
+ \mu_i' V^{-1}(\hat{V} - V)V^{-1}(\hat{V} - V)V^{-1}\mu_i + (\hat{\mu}_r - \mu_r)^T V^{-1}(\hat{\mu}_r - \mu_r)
\]

\[
(\hat{c}_{kk} - c_{kk})(\hat{c}_n - c_n)
\]

\[
\equiv \mu_k' V^{-1}(\hat{V} - V)V^{-1}\mu_k \mu_i' V^{-1}(\hat{V} - V)V^{-1}\mu_i + 4(\hat{\mu}_k - \mu_k)^T V^{-1} \mu_k \mu_i' V^{-1}(\hat{\mu}_r - \mu_r)
\]

\[
(\hat{a} - a)(\hat{c}_n - c_n) \equiv 1V^{-1}(\hat{V} - V)V^{-1} \mu_i' V^{-1}(\hat{V} - V)V^{-1}\mu_i
\]

\[
(\hat{b}_k - b_k)(\hat{c}_n - c_n) \equiv 1V^{-1}(\hat{V} - V)V^{-1} \mu_k \mu_i' V^{-1}(\hat{V} - V)V^{-1}\mu_i + 2(\hat{\mu}_k - \mu_k)^T V^{-1} \mu_i' V^{-1}(\hat{\mu}_r - \mu_r)
\]

\[
\hat{d}_n - d_n \equiv -a \mu_i' V^{-1}(\hat{V} - V)V^{-1}\mu_i + (-c_n 1 + 2b \mu_i)^T V^{-1}(\hat{V} - V)V^{-1} 1
\]

\[
+ 2(a \mu_i' - b_i 1)^T V^{-1}(\hat{\mu}_r - \mu_r) + a \mu_i' V^{-1}(\hat{V} - V)V^{-1}(\hat{V} - V)V^{-1}\mu_i
\]

\[
+ 1V^{-1}(\hat{V} - V)V^{-1}(\hat{V} - V)V^{-1}(c_n 1 - 2b \mu_i)
\]

\[
+ 1V^{-1}(\hat{V} - V)V^{-1}(1_4' - \mu_i')(\hat{V} - V)V^{-1}\mu_i + (\hat{\mu}_r - \mu_r)^T (a V^{-1} - V^{-1} 11 V^{-1})(\hat{\mu}_r - \mu_r)
\]

\[
(\hat{d}_{kk} - d_{kk})(\hat{d}_n - d_n) \equiv a^2 \mu_k' V^{-1}(\hat{V} - V)V^{-1}\mu_k \mu_i' V^{-1}(\hat{V} - V)V^{-1}\mu_i
\]

\[
- a \mu_k' V^{-1}(\hat{V} - V)V^{-1}\mu_k (-c_n 1 + 2b \mu_i)^T V^{-1}(\hat{V} - V)V^{-1} 1
\]

\[
- a(-c_{kk} 1 + 2b \mu_k)^T V^{-1}(\hat{V} - V)V^{-1} \mu_i' V^{-1}(\hat{V} - V)V^{-1}\mu_i
\]

\[
+ (a \mu_k' - b_i 1)^T (\hat{V} - V)V^{-1} (-c_n 1 + 2b \mu_i)^T (\hat{V} - V)V^{-1} 1
\]

\[
+ 4(\hat{\mu}_k - \mu_k)^T (a \mu_k' - b_i 1)(a \mu_i' - b_i 1)^T V^{-1}(\hat{\mu}_r - \mu_r)
\]
\[
(\hat{d}_a - d_a) = a^2 \mu'_V V^{-1}(\hat{V} - V)V^{-1} \mu'_V (\hat{V} - V)V^{-1} \mu_r \\
- a \mu'_V V^{-1}(\hat{V} - V)V^{-1} \mu_r (-c_a 1 + 2b \mu_r) V^{-1} (\hat{V} - V)V^{-1} 1 \\
- a(-c_a 1 + 2b \mu_r) V^{-1} (\hat{V} - V)V^{-1} \mu'_V (\hat{V} - V)V^{-1} \mu_r, \\
+ (-c_a 1 + 2b \mu_r) V^{-1} (\hat{V} - V)V^{-1} 1 (-c_a 1 + 2b \mu_r) V^{-1} (\hat{V} - V)V^{-1} 1 \\
+ 4(\hat{\mu}_r - \mu_r) V^{-1} (a \mu_r - b_r)^{'} V^{-1} (\hat{\mu}_r - \mu_r) \\

(\hat{a} - a)(\hat{d}_a - d_a) \\
\equiv a 1' V^{-1} (\hat{V} - V)V^{-1} \mu'_V (\hat{V} - V)V^{-1} \mu_r + 1' V^{-1} (\hat{V} - V)V^{-1} 1 (c_a 1 - 2b \mu_r) V^{-1} (\hat{V} - V)V^{-1} 1 \\

(\hat{b}_k - b_k)(\hat{d}_a - d_a) \equiv a 1' V^{-1} (\hat{V} - V)V^{-1} \mu'_V (\hat{V} - V)V^{-1} \mu_r \\
- 1' V^{-1} (\hat{V} - V)V^{-1} \mu_k (-c_a 1 + 2b \mu_r) V^{-1} (\hat{V} - V)V^{-1} 1 \\
+ 2(\hat{\mu}_k - \mu_k) V^{-1} 1 (a \mu_r - b_r)^{'} V^{-1} (\hat{\mu}_r - \mu_r) \\

(\hat{c}_{kk} - c_{kk})(\hat{d}_a - d_a) \equiv a \mu'_V V^{-1}(\hat{V} - V)V^{-1} \mu'_V (\hat{V} - V)V^{-1} \mu_r \\
- \mu'_V (\hat{V} - V)V^{-1} \mu_k (-c_a 1 + 2b \mu_r) V^{-1} (\hat{V} - V)V^{-1} 1 \\
+ 4(\hat{\mu}_k - \mu_k) V^{-1} \mu_k (a \mu_r - b_r)^{'} V^{-1} (\hat{\mu}_r - \mu_r) \\

(\hat{c}_a - c_a)(\hat{d}_a - d_a) \equiv a \mu'_V V^{-1}(\hat{V} - V)V^{-1} \mu'_V (\hat{V} - V)V^{-1} \mu_r \\
- \mu V^{-1} (\hat{V} - V)V^{-1} \mu_r (-c_a 1 + 2b \mu_r) V^{-1} (\hat{V} - V)V^{-1} 1 \\
+ 4(\hat{\mu}_r - \mu_r) V^{-1} \mu_r (a \mu_r - b_r)^{'} V^{-1} (\hat{\mu}_r - \mu_r) \\

(A.132)
Proof: Using the expansion \( \hat{V}^{-1} \approx V^{-1} - V^{-1}(\hat{V} - V)V^{-1} + V^{-1}(\hat{V} - V)V^{-1}(\hat{V} - V)V^{-1} \) from Lemma 4, we find

\[
\hat{a} - a = 1\hat{V}^{-1}1 - 1V^{-1}1 \approx 1\left[ V^{-1} - V^{-1}(\hat{V} - V)V^{-1} + V^{-1}(\hat{V} - V)V^{-1}(\hat{V} - V)V^{-1} \right]1 - 1V^{-1}1
\]

\[
= -1V^{-1}(\hat{V} - V)V^{-1}1 + 1V^{-1}(\hat{V} - V)V^{-1}(\hat{V} - V)V^{-1}1
\]

\[
(\hat{a} - a)^2 \approx \left[-1V^{-1}(\hat{V} - V)V^{-1}1 \right]^2 = 1V^{-1}(\hat{V} - V)V^{-1}11V^{-1}(\hat{V} - V)V^{-1}1
\]

\[
\hat{b}_r - b_r = 1\hat{V}^{-1}\hat{\mu}_r - 1V^{-1}\mu_r \equiv 1\left(\hat{V}^{-1} - V^{-1}\right)\mu_r + 1V^{-1}(\hat{\mu}_r - \mu_r)
\]

\[
\approx 1\left[-V^{-1}(\hat{V} - V)V^{-1} + V^{-1}(\hat{V} - V)V^{-1}(\hat{V} - V)V^{-1} \right]\mu_r + 1V^{-1}(\hat{\mu}_r - \mu_r)
\]

\[
= -1V^{-1}(\hat{V} - V)V^{-1}\mu_r + 1V^{-1}(\hat{\mu}_r - \mu_r) + 1V^{-1}(\hat{V} - V)V^{-1}(\hat{V} - V)V^{-1}\mu_r
\]

\[
(\hat{b}_r - b_r)(\hat{b}_r - b_r) \approx \left[-1V^{-1}(\hat{V} - V)V^{-1}1 \right]\left[-1V^{-1}(\hat{V} - V)V^{-1}\mu_r \right]
\]

\[
= 1V^{-1}(\hat{V} - V)V^{-1}11V^{-1}(\hat{V} - V)V^{-1}\mu_r
\]

\[
\hat{c}_r - c_r = \hat{\mu}_r\hat{V}^{-1}\hat{\mu}_r - \mu_rV^{-1}\mu_r
\]

\[
\approx \mu_r\left(\hat{V}^{-1} - V^{-1}\right)\mu_r + 2\mu_rV^{-1}(\hat{\mu}_r - \mu_r) + (\hat{\mu}_r - \mu_r)V^{-1}(\hat{\mu}_r - \mu_r)
\]

\[
= \mu_r\left[-V^{-1}(\hat{V} - V)V^{-1} + V^{-1}(\hat{V} - V)V^{-1}(\hat{V} - V)V^{-1} \right]\mu_r
\]

\[
+ 2\mu_rV^{-1}(\hat{\mu}_r - \mu_r) + (\hat{\mu}_r - \mu_r)V^{-1}(\hat{\mu}_r - \mu_r)
\]

\[
= -\mu_rV^{-1}(\hat{V} - V)V^{-1}\mu_r + 2\mu_rV^{-1}(\hat{\mu}_r - \mu_r)
\]

\[
+ \mu_rV^{-1}(\hat{V} - V)V^{-1}(\hat{V} - V)V^{-1}\mu_r + (\hat{\mu}_r - \mu_r)V^{-1}(\hat{\mu}_r - \mu_r)
\]

\[
= \mu_rV^{-1}(\hat{V} - V)V^{-1}\mu_r + (\hat{\mu}_r - \mu_r)V^{-1}(\hat{\mu}_r - \mu_r)
\]
\((\hat{c}_{ik} - c_{ik})(\hat{c}_{ii} - c_{ii})\)
\[
\cong \left[ -\mu_k'V^{-1}(\hat{V} - V)V^{-1}\mu_k + 2\mu_k'V^{-1}(\hat{\mu}_k - \mu_k) \right]\left[ -\mu_i'V^{-1}(\hat{V} - V)V^{-1}\mu_i + 2\mu_i'V^{-1}(\hat{\mu}_i - \mu_i) \right] \\
\cong \mu_k'V^{-1}(\hat{V} - V)V^{-1}\mu_k \mu_i'V^{-1}(\hat{V} - V)V^{-1}\mu_i + 4(\hat{\mu}_k - \mu_k)'V^{-1}\mu_k \mu_i'V^{-1}(\hat{\mu}_i - \mu_i) \\
\]

\((\hat{a} - a)(\hat{c}_{ii} - c_{ii})\)\[\cong \left[ -1V^{-1}(\hat{V} - V)V^{-1}1 \right]\left[ -\mu_i'V^{-1}(\hat{V} - V)V^{-1}\mu_i \right] \\
= 1V^{-1}(\hat{V} - V)V^{-1}1\mu_i'V^{-1}(\hat{V} - V)V^{-1}\mu_i \]

\((\hat{b}_k - b_k)(\hat{c}_{ii} - c_{ii})\)
\[
= \left[ -1V^{-1}(\hat{V} - V)V^{-1}\mu_k + 1V^{-1}(\hat{\mu}_k - \mu_k) \right]\left[ -\mu_i'V^{-1}(\hat{V} - V)V^{-1}\mu_i + 2\mu_i'V^{-1}(\hat{\mu}_i - \mu_i) \right] \\
\cong \left[ -1V^{-1}(\hat{V} - V)V^{-1}\mu_k \right]\left[ -\mu_i'V^{-1}(\hat{V} - V)V^{-1}\mu_i \right] + \left[ 1V^{-1}(\hat{\mu}_k - \mu_k) \right]\left[ 2\mu_i'V^{-1}(\hat{\mu}_i - \mu_i) \right] \\
= 1V^{-1}(\hat{V} - V)V^{-1}\mu_k \mu_i'V^{-1}(\hat{V} - V)V^{-1}\mu_i + 2(\hat{\mu}_k - \mu_k)'V^{-1}\mu_k \mu_i'V^{-1}(\hat{\mu}_i - \mu_i) \]
\begin{align*}
\dot{d}_n - d_n &= a(\dot{c}_n - c_n) + c_n (\dot{a} - a) + (\dot{a} - a)(\dot{c}_n - c_n) - 2b\left(\dot{b}_r - b_r\right) - \left(\dot{b}_r - b_r\right)^2 \\
&\approx a\left[-\mu'V^{-1}(\dot{V} - V)V^{-1}\mu_r + 2\mu'V^{-1}(\hat{\mu}_r - \mu_r)\right] \\
&+ a\left[\mu'V^{-1}(\dot{V} - V)V^{-1}(\dot{V} - V)V^{-1}\mu_r + (\hat{\mu}_r - \mu_r)'V^{-1}(\hat{\mu}_r - \mu_r)\right] \\
&+ c_n\left[-1V^{-1}(\dot{V} - V)V^{-1} + 1V^{-1}(\dot{V} - V)V^{-1}(\dot{V} - V)V^{-1}\right] \\
&+ [1V^{-1}(\dot{V} - V)V^{-1}]\left[-\mu'V^{-1}(\dot{V} - V)V^{-1}\mu_r + 2\mu'V^{-1}(\hat{\mu}_r - \mu_r)\right] \\
&- 2b\left[-1V^{-1}(\dot{V} - V)V^{-1}\mu_r + 1V^{-1}(\hat{\mu}_r - \mu_r) + 1V^{-1}(\dot{V} - V)V^{-1}(\dot{V} - V)V^{-1}\mu_r\right] \\
&- \left[-1V^{-1}(\dot{V} - V)V^{-1}\mu_r + 1V^{-1}(\hat{\mu}_r - \mu_r)\right]^2 \\
&\approx -a\mu'V^{-1}(\dot{V} - V)V^{-1}\mu_r + 2a\mu'V^{-1}(\hat{\mu}_r - \mu_r) \\
&+ a\mu'V^{-1}(\dot{V} - V)V^{-1}(\dot{V} - V)V^{-1}\mu_r + a(\hat{\mu}_r - \mu_r)'V^{-1}(\hat{\mu}_r - \mu_r) \\
&- c_n 1V^{-1}(\dot{V} - V)V^{-1} + c_n 1V^{-1}(\dot{V} - V)V^{-1}(\dot{V} - V)V^{-1} \\
&+ 1V^{-1}(\dot{V} - V)V^{-1}\mu_r 1V^{-1}(\dot{V} - V)V^{-1}\mu_r \\
&+ 2b\left[1V^{-1}(\dot{V} - V)V^{-1}\mu_r - 2b\left[1V^{-1}(\hat{\mu}_r - \mu_r) - 2b\left[1V^{-1}(\dot{V} - V)V^{-1}(\dot{V} - V)V^{-1}\mu_r\right] \\
&- 1V^{-1}(\dot{V} - V)V^{-1}\mu_r 1V^{-1}(\dot{V} - V)V^{-1}\mu_r - (\hat{\mu}_r - \mu_r)'V^{-1}11V^{-1}(\hat{\mu}_r - \mu_r)
\end{align*}

Continuing, we find

\begin{align*}
\hat{d}_n - d_n &= a(\hat{c}_n - c_n) + c_n (\hat{a} - a) + (\hat{a} - a)(\hat{c}_n - c_n) - 2b\left(\hat{b}_r - b_r\right) - \left(\hat{b}_r - b_r\right)^2 \\
&\approx a\left[-\mu'V^{-1}(\dot{V} - V)V^{-1}\mu_r + 2\mu'V^{-1}(\hat{\mu}_r - \mu_r)\right] \\
&+ a\left[\mu'V^{-1}(\dot{V} - V)V^{-1}(\dot{V} - V)V^{-1}\mu_r + (\hat{\mu}_r - \mu_r)'V^{-1}(\hat{\mu}_r - \mu_r)\right] \\
&+ c_n\left[-1V^{-1}(\dot{V} - V)V^{-1} + 1V^{-1}(\dot{V} - V)V^{-1}(\dot{V} - V)V^{-1}\right] \\
&+ [1V^{-1}(\dot{V} - V)V^{-1}]\left[-\mu'V^{-1}(\dot{V} - V)V^{-1}\mu_r + 2\mu'V^{-1}(\hat{\mu}_r - \mu_r)\right] \\
&- 2b\left[-1V^{-1}(\dot{V} - V)V^{-1}\mu_r + 1V^{-1}(\hat{\mu}_r - \mu_r) + 1V^{-1}(\dot{V} - V)V^{-1}(\dot{V} - V)V^{-1}\mu_r\right] \\
&- \left[-1V^{-1}(\dot{V} - V)V^{-1}\mu_r + 1V^{-1}(\hat{\mu}_r - \mu_r)\right]^2 \\
&\approx -a\mu'V^{-1}(\dot{V} - V)V^{-1}\mu_r + 2a\mu'V^{-1}(\hat{\mu}_r - \mu_r) \\
&+ a\mu'V^{-1}(\dot{V} - V)V^{-1}(\dot{V} - V)V^{-1}\mu_r + a(\hat{\mu}_r - \mu_r)'V^{-1}(\hat{\mu}_r - \mu_r) \\
&- c_n 1V^{-1}(\dot{V} - V)V^{-1} + c_n 1V^{-1}(\dot{V} - V)V^{-1}(\dot{V} - V)V^{-1} \\
&+ 1V^{-1}(\dot{V} - V)V^{-1}\mu_r 1V^{-1}(\dot{V} - V)V^{-1}\mu_r \\
&+ 2b\left[1V^{-1}(\dot{V} - V)V^{-1}\mu_r - 2b\left[1V^{-1}(\hat{\mu}_r - \mu_r) - 2b\left[1V^{-1}(\dot{V} - V)V^{-1}(\dot{V} - V)V^{-1}\mu_r\right] \\
&- 1V^{-1}(\dot{V} - V)V^{-1}\mu_r 1V^{-1}(\dot{V} - V)V^{-1}\mu_r - (\hat{\mu}_r - \mu_r)'V^{-1}11V^{-1}(\hat{\mu}_r - \mu_r)
\end{align*}
\((\hat{d}_{kk} - d_{kk})(\hat{d}_{n} - d_{n})\)
\[\approx \left[ -a\mu'_k V \left( \hat{V} - V \right) V^{-1}C_k + (-c_k \mathbf{1} + 2b_k \mu_k) V^{-1} \left( \hat{V} - V \right) V^{-1} \mathbf{1} + 2(a\mu' - b_k \mathbf{1}^V) V^{-1} (\hat{\mu}_k - \mu_k) \right] \]
\[\approx \left[ -a\mu'_k V \left( \hat{V} - V \right) V^{-1}C_k + (-c_k \mathbf{1} + 2b_k \mu_k) V^{-1} \left( \hat{V} - V \right) V^{-1} \mathbf{1} + 2(a\mu' - b_k \mathbf{1}^V) V^{-1} (\hat{\mu}_k - \mu_k) \right] \]
\[= a^2\mu'_k V \left( \hat{V} - V \right) V^{-1}C_k + (-c_k \mathbf{1} + 2b_k \mu_k) V^{-1} \left( \hat{V} - V \right) V^{-1} \mathbf{1} + a(-c_k \mathbf{1} + 2b_k \mu_k) V^{-1} \left( \hat{V} - V \right) V^{-1} \mathbf{1} + 4(\hat{\mu}_k - \mu_k) \left( a\mu' - b_k \mathbf{1} \right) V^{-1} (\hat{\mu}_k - \mu_k) \]

\((\hat{a} - a)(\hat{d}_{n} - d_{n})\)
\[\approx \left[ -1 \left( \hat{V} - V \right) V^{-1} \mathbf{1} \right] \left[ -a\mu'_k V \left( \hat{V} - V \right) V^{-1}C_k + (-c_k + 2b_k \mu'_k) V^{-1} (\hat{V} - V) V^{-1} \mathbf{1} \right] \]
\[= a \left( \hat{V} - V \right) V^{-1} C_k \left( \hat{V} - V \right) V^{-1} \mathbf{1} + \mathbf{1} \left( \hat{V} - V \right) V^{-1} (c_k \mathbf{1} + 2b \mu_k) \left( \hat{V} - V \right) V^{-1} \mathbf{1} \]
\[
(\hat{b}_k - b_k)(\hat{d}_n - d_n)
\]
\[
\cong \left[ -1 V^{-1}(\hat{V} - V)V^{-1}\mu_k + I V^{-1}(\hat{\mu}_k - \mu_k) \right]
\]
\[
\left[ -a\mu'V^{-1}(\hat{V} - V)V^{-1}\mu_k + (-c_n + 2b\eta) V^{-1}(\hat{V} - V)V^{-1}1 + 2(a\mu' - b_1')V^{-1}(\hat{\mu}_k - \mu_k) \right]
\]
\[
\cong -1 V^{-1}(\hat{V} - V)V^{-1}\mu_k \left[ -a\mu'V^{-1}(\hat{V} - V)V^{-1}\mu_k + (-c_n + 2b\eta) V^{-1}(\hat{V} - V)V^{-1}1 + 2(a\mu' - b_1')V^{-1}(\hat{\mu}_k - \mu_k) \right] + 2(\hat{\mu}_k - \mu_k)' V^{-1}1 (a\mu' - b_1') V^{-1}(\hat{\mu}_k - \mu_k)
\]
\[
= a1 V^{-1}(\hat{V} - V)V^{-1}\mu_k V^{-1}(\hat{V} - V)V^{-1}\mu_k
\]
\[
- 1 V^{-1}(\hat{V} - V)V^{-1}\mu_k (-c_n + 2b\eta) V^{-1}(\hat{V} - V)V^{-1}1
\]
\[
+ 2(\hat{\mu}_k - \mu_k)' V^{-1}1 (a\mu' - b_1') V^{-1}(\hat{\mu}_k - \mu_k)
\]

The expression for \((\hat{b}_j - b_j)(\hat{d}_n - d_n)\) follows directly.

\[
(\hat{c}_{kk} - c_{kk})(\hat{d}_n - d_n)
\]
\[
\cong \left[ -\mu'V^{-1}(\hat{V} - V)V^{-1}\mu_k + 2\mu'V^{-1}(\hat{\mu}_k - \mu_k) \right]
\]
\[
\left[ -a\mu'V^{-1}(\hat{V} - V)V^{-1}\mu_k + (-c_n + 2b\eta) V^{-1}(\hat{V} - V)V^{-1}1 + 2(a\mu' - b_1')V^{-1}(\hat{\mu}_k - \mu_k) \right] + 2(\hat{\mu}_k - \mu_k)' V^{-1}1 (a\mu' - b_1') V^{-1}(\hat{\mu}_k - \mu_k)
\]
\[
\cong a\mu'V^{-1}(\hat{V} - V)V^{-1}\mu_k V^{-1}(\hat{V} - V)V^{-1}\mu_k
\]
\[
- \mu'V^{-1}(\hat{V} - V)V^{-1}\mu_k (-c_n + 2b\eta) V^{-1}(\hat{V} - V)V^{-1}1
\]
\[
+ 4(\hat{\mu}_k - \mu_k)' V^{-1}\mu_k (a\mu' - b_1) V^{-1}(\hat{\mu}_k - \mu_k)
\]

The expression for \((\hat{c}_n - c_n)(\hat{d}_n - d_n)\) follows directly, completing the proof. □

**Lemma 9:** Expectations are as follows:

\[
E(\hat{a} - a) \cong \frac{a(n + 1)}{T - L}
\]

\[
E\left[(\hat{a} - a)^2\right] \cong \frac{2a^2}{T - L}
\]
\[
E\left( \hat{b}_i - b_i \right) \approx \frac{(n+1)}{T - L} b_i
\]

\[
E\left[ (\hat{b}_k - b_k)(\hat{b}_i - b_i) \right] \approx \frac{a(c_{ikt} - 1) + b_k b_i}{T - L} + \frac{a}{T} Z'_{i-1} A^{-1} Z_{i-1}
\]

\[
E\left[ (\hat{a} - a)(\hat{b}_i - b_i) \right] \approx \frac{2ab_i}{T - L}
\]

\[
E\left( \hat{c}_n - c_n \right) \approx \frac{(n+1)(c_n - 1)}{T - L} + \frac{n}{T} Z'_{i-1} A^{-1} Z_{i-1}
\]

\[
E\left[ (\hat{c}_{kk} - c_{kk})(\hat{c}_n - c_n) \right] \approx \frac{2(c_{kk} - 1)^2}{T - L} + \frac{4(c_{kk} - 1)}{T} Z'_{i-1} A^{-1} Z_{i-1}
\]

\[
E\left[ (\hat{c}_n - c_n) \right] \approx \frac{2b^2_i}{T - L}
\]

\[
E\left[ (\hat{b}_k - b_k)(\hat{c}_n - c_n) \right] \approx \frac{2b_k (c_{ikt} - 1)}{T - L} + \frac{2b_k}{T} Z'_{i-1} A^{-1} Z_{i-1}
\]

\[
E\left( \hat{d}_{it} - d_{it} \right) \approx \frac{-na + (2n+1)d_{it}}{T - L} + \frac{a(n-1)}{T} Z'_{i-1} A^{-1} Z_{i-1}
\]

\[
E\left[ (\hat{d}_{kk} - d_{kk})(\hat{d}_n - d_n) \right] \approx \frac{-2ab_k b c_{ikt} + 2a^2 + 2a^2 c_{kk} c_{nt} - 2b_k^2 d_{it} - 2b_k^2 d_{kk} - 4a d_{ik} + 2a c_{ikt} d_{ikt} + 4 a d_{ikt} - a^2}{T - L} Z'_{i-1} A^{-1} Z_{i-1}
\]

\[
E\left[ (\hat{d}_n - d_n)^2 \right] \approx \frac{2a^2 - 4a d_{it} + 4d^2_{it} + 4 a d_{it} - a^2}{T - L} Z'_{i-1} A^{-1} Z_{i-1}
\]
\[
\frac{E(\hat{\mu}_n - \mu_n)^2}{d_n^2} - \frac{E(\hat{\mu}_n - \mu_n)}{T - L}
\]
\[
= \frac{2a^2 + (n-4)d_n + (3-2n)d_n^2}{(T-L)d_n^3} + \frac{-4a^2 - (n-5)d_n}{Td_n^3} Z_{t-1}^{-1} A^{-1} Z_{t-1}
\]

\[
E[\hat{\alpha} - a)(\hat{\mu}_n - \mu_n)] \approx \frac{2ad_n}{T - L}
\]

\[
E[(\hat{b}_k - b_k)(\hat{\mu}_n - \mu_n)] \approx \frac{2b_k d_n}{T - L}
\]

\[
E[(\hat{b}_k - b_k)(\hat{\mu}_n - \mu_n)] \approx \frac{2b_k d_n}{T - L}
\]

\[
E[(\hat{c}_{kk} - c_{kk})(\hat{\mu}_n - \mu_n)] \approx \frac{2a(c_{kk} - 1)^2 - 4b_k b_{ik}(c_{k} - 1) + 2b_k^2 c_{ik} + 4d_{ik} - a}{T} Z_{k-1}^{-1} A^{-1} Z_{t-1}
\]

\[
E[(\hat{c}_n - c_n)(\hat{\mu}_n - \mu_n)] \approx \frac{2a(c_{ik} - 1)^2 - 2b_k(c_{i} - 2) + 4d_{ik} - a}{T} Z_{k-1}^{-1} A^{-1} Z_{t-1}
\]  
(A.136)

**Proof:** Making use of unbiasedness of $\hat{\mu}_i$ and of $\hat{V}$, along with Lemmas 5 and 6, and commutativity within the trace operator, together with the expansions from Lemma 8, we find

\[
E(\hat{\alpha} - a) = 1' V^{-1} E[(\hat{V} - V)V^{-1}(\hat{V} - V)]V^{-1} 1 = 1' V^{-1} \frac{(n+1)V}{T-L} V^{-1} 1 = \frac{a(n+1)}{T-L}
\]

\[
E[(\hat{\alpha} - a)^2] = E[1' V^{-1} (\hat{V} - V)V^{-1} 1 1' V^{-1} (\hat{V} - V)V^{-1} 1] = \frac{2a^2}{T-L}
\]

\[
E(\hat{b}_i - b_i) = 1' V^{-1} E[(\hat{V} - V)V^{-1}(\hat{V} - V)]V^{-1} \mu_i = 1' V^{-1} \frac{(n+1)V}{T-L} V^{-1} \mu_i = \frac{(n+1)b_i}{T-L}
\]
\[
E \left[ (\hat{b}_k - b_k)(\hat{b}_i - b_i) \right] \\
\approx E \left[ 1'V^{-1}(\hat{V} - V)V^{-1}\mu_k' 1'V^{-1}(\hat{V} - V)V^{-1}\mu_i \right] + E \left[ (\hat{\mu}_k - \mu_k)' V^{-1} 11'V^{-1} (\hat{\mu}_i - \mu_i) \right] \\
= \frac{a(c_{k\ell} - 1) + b_kb_i}{T - L} + \frac{a}{T} Z'_{t-1} A^{-1} Z_{t-1}
\]

\[
E \left[ (\hat{a} - a)(\hat{b}_k - b_k) \right] \approx E \left[ 1'V^{-1}(\hat{V} - V)V^{-1} 11'V^{-1}(\hat{V} - V)V^{-1}\mu_i \right] = \frac{2ab_i}{T - L}
\]

\[
E \left[ (\hat{c}_n - c_n) \right] \approx \mu_i'V^{-1}E \left[ (\hat{V} - V)V^{-1}(\hat{V} - V) \right] V^{-1}\mu_i + E \left[ (\hat{\mu}_n - \mu_n)' V^{-1} (\hat{\mu}_n - \mu_n) \right] \\
= \mu_i' V^{-1} \left( n + 1 \right) V^{-1} \mu_i + \frac{n}{T} Z'_{t-1} A^{-1} Z_{t-1} = \frac{(n + 1) (c_n - 1)}{T - L} + \frac{n}{T} Z'_{t-1} A^{-1} Z_{t-1}
\]

\[
E \left[ (\hat{c}_{kk} - c_{kk})(\hat{c}_n - c_n) \right] \\
\approx E \left[ \mu_k' V^{-1} (\hat{V} - V) V^{-1} \mu_k' V^{-1} (\hat{V} - V) V^{-1}\mu_i \right] + 4E \left[ (\hat{\mu}_k - \mu_k)' V^{-1} \mu_k' V^{-1} (\hat{\mu}_n - \mu_n) \right] \\
= \frac{2(c_{k\ell} - 1)^2}{T - L} + \frac{4\mu_k' V^{-1} \mu_k Z'_{t-1} A^{-1} Z_{t-1}}{T - L} = \frac{2(c_{k\ell} - 1)^2}{T - L} + \frac{4(c_{k\ell} - 1)}{T} Z'_{t-1} A^{-1} Z_{t-1}
\]

\[
E \left[ (\hat{a} - a)(\hat{c}_n - c_n) \right] \approx E \left[ 1'V^{-1}(\hat{V} - V)V^{-1} 1\mu_i' V^{-1}(\hat{V} - V)V^{-1}\mu_i \right] \\
= \frac{2b_i^2}{T - L}
\]

\[
E \left[ (\hat{b}_k - b_k)(\hat{c}_n - c_n) \right] \\
\approx E \left[ 1'V^{-1}(\hat{V} - V)V^{-1} \mu_k' V^{-1}(\hat{V} - V)V^{-1}\mu_i \right] + 2E \left[ (\hat{\mu}_k - \mu_k)' V^{-1} 1\mu_i' V^{-1}(\hat{\mu}_n - \mu_n) \right] \\
= \frac{1'V^{-1} \mu_k' V^{-1} \mu_i + 1'V^{-1} \mu_i' V^{-1} \mu_i}{T - L} + 2\frac{1'V^{-1} \mu_i Z'_{t-1} A^{-1} Z_{t-1}}{T - L} = \frac{2b_i (c_{k\ell} - 1)}{T - L} + \frac{2b_i}{T} Z'_{t-1} A^{-1} Z_{t-1}
\]
\[
E(\hat{d}_n - d_n) \equiv a\mu V^{-1}E[(\hat{V} - V)V^{-1}(\hat{V} - V)]V^{-1}\mu,
\]
\[+ \mathbf{1}V^{-1}E[(\hat{V} - V)V^{-1}(\hat{V} - V)]V^{-1}(c_n \mathbf{1} - 2b\mu),
\]
\[+ E[\mathbf{1}V^{-1}(\hat{V} - V)V^{-1}(\mathbf{1} \mu' - \mu, \mathbf{1}')V^{-1}(\hat{V} - V)V^{-1}\mu],
\]
\[+ E[(\hat{\mu'} - \mu')(aV^{-1} - V^{-1} \mathbf{1}\mathbf{1}V^{-1})(\hat{\mu'} - \mu')]
\]
\[= a\mu V^{-1}\frac{(n+1)V}{T - L} - \mu + \mathbf{1}V^{-1}\frac{(n+1)V}{T - L}V^{-1}(c_n \mathbf{1} - 2b\mu),
\]
\[+ \frac{2b^2}{T - L} - \frac{T - L}{T - L} + \frac{an}{T} Z_{t-1}^\prime A^{-1}Z_{t-1} - \frac{1V^{-1}\mathbf{1}}{T} Z_{t-1}^\prime A^{-1}Z_{t-1}
\]
\[= a\left(c_n - 1\right)(n+1) - \frac{ac_n - b^2}{T - L} + \frac{an}{T} Z_{t-1}^\prime A^{-1}Z_{t-1}
\]
\[= -\frac{a(2ac_n - 2b^2)(n+1)}{T - L} + \frac{an}{T} Z_{t-1}^\prime A^{-1}Z_{t-1}
\]
\[= \frac{-a + 2d_n}{T - L} + \frac{an}{T} Z_{t-1}^\prime A^{-1}Z_{t-1}
\]
\[= \frac{-na + (2n+1)d_n}{T - L} + \frac{an}{T} Z_{t-1}^\prime A^{-1}Z_{t-1}
\]
\[
E \left[ (\hat{d}_{kk} - d_{kk})(\hat{d}_{nn} - d_{nn}) \right] \approx a^2 E \left[ \mu_k'V^{-1}(\hat{V} - V)V^{-1}\mu_k'V^{-1}(\hat{V} - V)V^{-1}\mu_r \right] \\
- aE \left[ \mu_k'V^{-1}(\hat{V} - V)V^{-1}\mu_k(-c_n^t + 2\bar{b}_k\mu_r)'V^{-1}(\hat{V} - V)V^{-1}\mathbf{1} \right] \\
- aE \left[ (-c_{kk}^t + 2\bar{b}_k\mu_r)'V^{-1}(\hat{V} - V)V^{-1}\mu_k'V^{-1}(\hat{V} - V)V^{-1}\mu_r \right] \\
+ E \left[ (-c_{kk}^t + 2\bar{b}_k\mu_r)'V^{-1}(\hat{V} - V)V^{-1}\mathbf{1}(-c_n^t + 2\bar{b}_k\mu_r)'V^{-1}(\hat{V} - V)V^{-1}\mathbf{1} \right] \\
+ 4E \left[ (\hat{\mu}_k - \mu_k)'V^{-1}(a\mu_k - \bar{b}_k\mathbf{1})(a\mu_k - \bar{b}_k\mathbf{1})'V^{-1}(\hat{\mu}_r - \mu_r) \right]
\]

\[
= a^2 \frac{2(c_{kk} - 1)^2}{T - L} + ac_{kk} E \left[ \mu_k'V^{-1}(\hat{V} - V)V^{-1}\mu_k'V^{-1}(\hat{V} - V)V^{-1}\mathbf{1} \right] \\
- 2ab_k E \left[ \mu_k'V^{-1}(\hat{V} - V)V^{-1}\mu_k'V^{-1}(\hat{V} - V)V^{-1}\mathbf{1} \right] \\
+ ac_{kk} E \left[ \mathbf{1}'V^{-1}(\hat{V} - V)V^{-1}\mu_k'V^{-1}(\hat{V} - V)V^{-1}\mu_r \right] \\
- 2ab_k E \left[ \mu_k'V^{-1}(\hat{V} - V)V^{-1}\mu_k'V^{-1}(\hat{V} - V)V^{-1}\mu_r \right] \\
+ c_{kk} c_n E \left[ \mathbf{1}'V^{-1}(\hat{V} - V)V^{-1}\mathbf{11}'V^{-1}(\hat{V} - V)V^{-1}\mathbf{1} \right] \\
- 2b_{kk} E \left[ \mathbf{1}'V^{-1}(\hat{V} - V)V^{-1}\mathbf{11}'V^{-1}(\hat{V} - V)V^{-1}\mathbf{1} \right] \\
- 2b_{kk} E \left[ \mu_k'V^{-1}(\hat{V} - V)V^{-1}\mathbf{11}'V^{-1}(\hat{V} - V)V^{-1}\mathbf{1} \right] \\
+ 4b_k E \left[ \mu_k'V^{-1}(\hat{V} - V)V^{-1}\mu_k'V^{-1}(\hat{V} - V)V^{-1}\mathbf{1} \right] \\
+ 4a^2 E \left[ (\hat{\mu}_k - \mu_k)'V^{-1}\mu_k'V^{-1}(\hat{\mu}_r - \mu_r) \right] - 4ab_k E \left[ (\hat{\mu}_k - \mu_k)'V^{-1}\mu_k'V^{-1}(\hat{\mu}_r - \mu_r) \right] \\
- 4ab_k E \left[ (\hat{\mu}_k - \mu_k)'V^{-1}\mu_k'V^{-1}(\hat{\mu}_r - \mu_r) \right] + 4b_k E \left[ (\hat{\mu}_k - \mu_k)'V^{-1}\mathbf{11}'V^{-1}(\hat{\mu}_r - \mu_r) \right]
\]

(A.137)

continuing, we find
\[ E\left[ (\hat{d}_{kk} - d_{kk}) (\hat{d}_{n} - d_{n}) \right] \approx a^2 \frac{2(c_{k_t}^2 - 2c_{k_t} + 1)}{T - L} + \frac{ac_{k}^2}{T - L} \frac{2b_k^2}{T - L} - 2ab_k \frac{2b_k c_{k_t} - 2b_k}{T - L} \\
+ \frac{ac_{k}^2}{T - L} - 2ab_k \frac{2b_k c_{k_t} - 2b_k}{T - L} + c_{k} c_{k_t} \frac{2a^2}{T - L} - 2b_k c_{k_t} \frac{2a^2}{T - L} - 2b_k c_{k_t} \frac{2a^2}{T - L} \\
+ 4b_k b_i \frac{ac_{k_t} - a + b_i b_i}{T - L} + 4a^2 \frac{(c_{k_t} - 1)}{T} Z_{t-1}^\prime A^{-1} Z_{k-1} - 4ab_k \frac{b_k}{T} Z_{t-1}^\prime A^{-1} Z_{k-1} \\
- 4ab_k \frac{b_i}{T} Z_{t-1}^\prime A^{-1} Z_{k-1} + 4b_k b_i \frac{a}{T} Z_{t-1}^\prime A^{-1} Z_{k-1} \\
= \frac{2a^2 c_{k_t}^2 - 4a^2 c_{k_t} + 2a^2 + 2ab_k^2 c_{k_t} - 4ab_k b_k}{T - L} \\
+ \frac{2ab_k^2 c_{k_t}}{T - L} - \frac{4ab_k b_i c_{k_t}}{T - L} - \frac{4ab_k b_i}{T - L} + \frac{2a^2 c_{k_t}^2 - 4a^2}{T} Z_{t-1}^\prime A^{-1} Z_{k-1} - \frac{4ab_k b_i}{T} Z_{t-1}^\prime A^{-1} Z_{k-1} \\
- \frac{4ab_k b_i}{T} Z_{t-1}^\prime A^{-1} Z_{k-1} + 4ab_k b_i \frac{a}{T} Z_{t-1}^\prime A^{-1} Z_{k-1} \\
= \frac{2a^2 + 4ab_k b_i - 4ab_k b_i + 4b_k^2 b_i^2 + 2ab_k^2 c_{k_t} - 4ab_k^2 c_{k_t}}{T - L} \\
+ \frac{4ab_k b_i c_{k_t} - 4ab_k b_i c_{k_t} - 4ab_k b_i c_{k_t} - 4ab_k^2 c_{k_t} + 2ab_k^2 c_{k_t} - 4ab_k^2 c_{k_t} + 2a^2 c_{k_t}^2 + 2a^2 c_{k_t} c_{k_t}}{T - L} \\
+ \frac{4 a^2 c_{k_t} - a^2 - ab_k b_i}{T} Z_{t-1}^\prime A^{-1} Z_{k-1} \\
= \frac{2a^2 + 4ab_k b_i + 4b_k^2 b_i^2 - 2ab_k^2 c_{k_t} - 4ab_k b_i c_{k_t} - 4a^2 c_{k_t} + 2a^2 c_{k_t}^2 + 2a^2 c_{k_t} c_{k_t}}{T - L} \\
+ \frac{4 a^2 c_{k_t} - a^2 - ab_k b_i}{T} Z_{t-1}^\prime A^{-1} Z_{k-1} \\
\text{continuing, we find} \]
\begin{align*}
E\left[ \left( \hat{d}_{kk} - d_{kk} \right) \left( \hat{d}_n - d_n \right) \right] &\approx \frac{-2ab_{kk} b_{c_{kk}} + 2a^2 + 2a^2 c_{kk} c_n - 2b_{kk}^2 \left( ac_n - b_t^2 \right)}{T - L} \\
&+ \frac{-2b_t^2 \left( ac_{kk} - b_k^2 \right) - 4a \left( ac_{kk} - b_k b_t \right) + 2a c_n \left( ac_{kk} - b_k b_t \right)}{T - L} \\
&+ \frac{4a \left( ac_{kk} - b_k b_t \right) - a^2}{T} \frac{Z_{t-1} A^{-1} Z_{k-1}}{T - L} \\
&= \frac{-2ab_{kk} b_{c_{kk}} + 2a^2 + 2a^2 c_{kk} c_n - 2b_{kk}^2 d_n - 2b_{kk}^2 d_{kk} - 4a d_n + 2a c_n d_n}{T - L} \\
&+ \frac{4a d_n - a^2}{T} \frac{Z_{t-1} A^{-1} Z_{k-1}}{T - L} \\
\end{align*}

\begin{align*}
E\left[ \left( \hat{d}_n - d_n \right)^2 \right] &\approx \frac{-2ab_{kk} b_{c_{kk}} + 2a^2 + 2a^2 c_{kk}^2 - 4b_t^2 d_n - 4a d_n + 2a c_n d_n}{T - L} \\
&+ \frac{2a c_n \left( ac_n - b_t^2 \right) + 2a^2 - 2b_t^2 d_n - 4a d_n + 2d_n \left( ac_n - b_t^2 \right)}{T - L} \\
&+ \frac{4a d_n - a^2}{T} \frac{Z_{t-1} A^{-1} Z_{t-1}}{T - L} \\
&= \frac{2a^2 + 2d_n \left( ac_n - b_t^2 \right) - 4a d_n + 2d_n^2}{T - L} \\
&+ \frac{4a d_n - a^2}{T} \frac{Z_{t-1} A^{-1} Z_{t-1}}{T - L} \\
&= \frac{2a^2 - 4a d_n + 4d_n^2}{T - L} + \frac{4a d_n - a^2}{T} \frac{Z_{t-1} A^{-1} Z_{t-1}}{T - L}
\end{align*}
\[
E\left(\hat{d}_n - d_n\right)^2 = \frac{E\left(\hat{d}_n - d_n\right)}{d_n^3} = \frac{E\left(\hat{d}_n - d_n\right)}{d_n^2} \\
\approx \frac{2a^2 - 4ad_n + 4d_n^2}{(T - L)d_n^3} + \frac{4ad_n - a^2}{Td_n^3} - Z_{t-1}Z_{t-1}^{-1}Z_{t-1}^{-1} - \frac{(n+1)ad_n}{Td_n^2}Z_{t-1}Z_{t-1}^{-1}Z_{t-1}^{-1} \\
= \frac{2a^2 - 4ad_n + 4d_n^2}{(T - L)d_n^3} - \frac{n(ad_n) + (2n + 1)d_n^2}{(T - L)d_n^3} + \frac{4ad_n - a^2}{Td_n^3} - Z_{t-1}Z_{t-1}^{-1}Z_{t-1}^{-1} \\
= \frac{2a^2 - 4ad_n + 4d_n^2}{(T - L)d_n^3} + \frac{4ad_n - a^2}{Td_n^3} - (n+1)ad_n Z_{t-1}Z_{t-1}^{-1}Z_{t-1}^{-1} \\
= \frac{2a^2 + (n-4)ad_n + (3-2n)d_n^2}{(T - L)d_n^3} + \frac{4ad_n - a^2}{Td_n^3} - (n-5)ad_n Z_{t-1}Z_{t-1}^{-1}Z_{t-1}^{-1} \\
= \frac{2a^2}{(T - L)} + \frac{2ad_n}{T - L} = \frac{2ad_n}{T - L} \\
\]
\[ E\left[ (\hat{b}_i - b_i)(\hat{d}_n - d_n)\right] \approx \frac{2b_d}{T-L} \]
then follows directly

\[ E\left[ (\hat{c}_{ik} - c_{ik})(\hat{d}_n - d_n)\right] \approx aE\left[ \mu_k'V^{-1}(\hat{V} - V)V^{-1}\mu_k'V^{-1}(\hat{V} - V)V^{-1}\mu_i\right] \\
- E\left[ \mu_k'V^{-1}(\hat{V} - V)V^{-1}\mu_k(-c_n1 + 2b_t\mu_t)'V^{-1}(\hat{V} - V)V^{-1}1\right] \\
+ 4E\left[ (\tilde{\mu}_k - \mu_k)'V^{-1}\mu_k(a\mu_t - b_t1)'V^{-1}(\tilde{\mu}_t - \mu_t)\right] \\
= \frac{2a(c_{ik} - 1)^2 + 2c_kb^2_k - 2b_t(c_{ki} - 1)b_k - 2b_t\mu_t(c_{ki} - 1)}{T-L} + \frac{4(a(c_{ki} - 1) - b_kb_t)}{T}Z'_{t-1}A^{-1}Z_{t-1} \\
= \frac{2a(c_{ik} - 1)^2 - 4b_kb_t(c_{ki} - 1) + 2b_t^2c_{nt}}{T-L} + \frac{4(d_n - a)}{T}Z'_{t-1}A^{-1}Z_{t-1} \\
\]

\[ E\left[ (\hat{c}_n - c_n)(\hat{d}_n - d_n)\right] \approx \frac{2a(c_n - 1)^2 - 4b_t^2(c_n - 1) + 2b^n_t c_{nt}}{T-L} + \frac{4(d_n - a)}{T}Z'_{t-1}A^{-1}Z_{t-1} \\
= \frac{2a(c_n - 1)^2 - 2b_t^2(c_n - 2)}{T-L} + \frac{4(d_n - a)}{T}Z'_{t-1}A^{-1}Z_{t-1} \\
\]

completing the proof. \( \Box \)

**Proof of Corollary I:**

We begin by noting that the portfolio return at time \( t \) is \( R_{pt} = w'R_n \) where \( R_t = \mu_t + \varepsilon_t = \delta'Z_{t-1} + \varepsilon_t \).

The estimated squared Sharpe Ratio is based on the estimated portfolio mean and the estimated portfolio variance. The conditional mean of \( R_{pt} \) is \( E\left( R_{pt} \mid Z_{t-1}\right) = w'\mu_r \), and the conditional variance of \( R_{pt} \) is \( \sigma_{p,t}^2 = \text{Var}\left( R_{pt} \mid Z_{t-1}\right) = w'Vw \) using the conditional covariance matrix \( V \) of \( R_t \) given \( Z_{t-1} \). The unconditional variance \( \sigma_p^2 \) of the portfolio \( R_p \) may then be expressed as the expected conditional variance plus the variance of the conditional portfolio means \( w'\mu_t \). Given a zero-beta rate \( \varphi \), the maximized squared-Sharpe-Ratio
\[ S_{\phi}^2 = (\bar{\mu} - \phi \mathbf{1})' U^{-1} (\bar{\mu} - \phi \mathbf{1}), \quad (A.140) \]

is constructed from the weights

\[ w_{\phi} = \frac{U^{-1} (\bar{\mu} - \phi \mathbf{1})}{\mathbf{1}' U^{-1} (\bar{\mu} - \phi \mathbf{1})} \quad (A.141) \]

for which the portfolio mean is \( \mu_{\phi} = \frac{\mu' U^{-1} (\bar{\mu} - \phi \mathbf{1})}{\mathbf{1}' U^{-1} (\bar{\mu} - \phi \mathbf{1})} \) and the portfolio variance is

\[ \sigma_{\phi}^2 = \frac{(\bar{\mu} - \phi \mathbf{1})' U^{-1} (\bar{\mu} - \phi \mathbf{1})}{\left[ \mathbf{1}' U^{-1} (\bar{\mu} - \phi \mathbf{1}) \right]^2} = \frac{S_{\phi}^2}{\left[ \mathbf{1}' U^{-1} (\bar{\mu} - \phi \mathbf{1}) \right]^2} \]

with \( S_{\phi}^2 = \frac{(\mu_{\phi} - \phi)^2}{\sigma_{\phi}^2} \), allowing us to write

\[ w_{\phi} = \frac{\sigma_{\phi}^2 U^{-1} (\bar{\mu} - \phi \mathbf{1})}{\mu_{\phi} - \phi}. \]

We begin with \( \hat{U}^{-1} = \left[ \hat{V} + \frac{1}{T} \sum_{t=1}^{T} \left( \hat{\mu}_t - \bar{\mu} \right) (\hat{\mu}_t - \bar{\mu})' \right]^{-1} \), which may be written as

\[ \hat{U}^{-1} = \left[ \hat{V} + \frac{1}{T} \sum_{t=1}^{T} \left( \hat{\mu}_t - \bar{\mu} \right) (\hat{\mu}_t - \bar{\mu})' \right]^{-1} + O_p \left( 1/T \right) \]

because \( (\hat{\mu} - \bar{\mu}) = O_p \left( 1/\sqrt{T} \right) \), so we have

\[ \frac{1}{T} \sum_{t=1}^{T} \left( \hat{\mu}_t - \bar{\mu} \right) (\hat{\mu}_t - \bar{\mu})' = \frac{1}{T} \sum_{t=1}^{T} (\hat{\mu}_t - \bar{\mu}) (\hat{\mu}_t - \bar{\mu})' + O_p \left( 1/T \right). \]

We expand as follows:

\[ \hat{U}^{-1} = \left[ \hat{V} + \frac{1}{T} \sum_{t=1}^{T} (\hat{\mu}_t - \bar{\mu}) (\hat{\mu}_t - \bar{\mu})' \right]^{-1} + O_p \left( 1/T \right) \]

\[ = U^{-1} - U^{-1} \left[ \hat{V} + \frac{1}{T} \sum_{t=1}^{T} (\hat{\mu}_t - \bar{\mu}) (\hat{\mu}_t - \bar{\mu})' - U \right] U^{-1} + O_p \left( 1/T \right) \]

\[ = U^{-1} - U^{-1} \left[ \hat{V} - V + \frac{1}{T} \sum_{t=1}^{T} \left( (\hat{\mu}_t - \mu_t)(\hat{\mu}_t - \mu_t)' + (\mu_t - \bar{\mu})(\hat{\mu}_t - \bar{\mu})' \right) \right] U^{-1} + O_p \left( 1/T \right) \]

Next, we expand the estimated squared Sharpe Ratio:
\[ \hat{S}_\phi^2 = (\hat{\mu} - \varphi \mathbf{1})' \hat{U}^{-1} (\hat{\mu} - \varphi \mathbf{1}) \]
\[ = S^2_\phi + 2 (\bar{\mu} - \varphi \mathbf{1})' U^{-1} (\hat{\mu} - \bar{\mu}) \]
\[ - (\bar{\mu} - \varphi \mathbf{1})' U^{-1} \left[ \bar{V} - \frac{1}{T} \sum_{t=1}^{T} \left[ (\hat{\mu}_t - \mu_t)(\mu_t - \bar{\mu})' + (\mu_t - \bar{\mu})(\hat{\mu}_t - \mu_t)' \right] \right] U^{-1} (\bar{\mu} - \varphi \mathbf{1}) \]
\[ + O_p \left( \frac{1}{T} \right) \]
\[ = S^2_\phi + 2 \frac{1}{T} \sum_{t=1}^{T} \left[ 1 - (\mu_t - \bar{\mu})' U^{-1} (\bar{\mu} - \varphi \mathbf{1}) \right] (\bar{\mu} - \varphi \mathbf{1})' U^{-1} (\hat{\mu}_t - \mu_t) \]
\[ - (\bar{\mu} - \varphi \mathbf{1})' U^{-1} \left( \bar{V} - \frac{1}{T} \sum_{t=1}^{T} (\mu_t - \bar{\mu})(\hat{\mu}_t - \mu_t)' \right) U^{-1} (\bar{\mu} - \varphi \mathbf{1}) + O_p \left( \frac{1}{T} \right) \]

from which we see that estimates of the canonical matrices are

\[ C = A^{-1} \frac{2}{T} \sum_{t=1}^{T} Z_{t-1} \left[ 1 - (\mu_t - \bar{\mu})' U^{-1} (\bar{\mu} - \varphi \mathbf{1}) \right] (\bar{\mu} - \varphi \mathbf{1})' U^{-1} \]  
\[ D = -U^{-1} (\bar{\mu} - \varphi \mathbf{1})(\bar{\mu} - \varphi \mathbf{1})' U^{-1} \]  

(A.144)

Using the fact that \( w = \frac{\sigma^2 \sigma^{-1} (\bar{\mu} - \varphi \mathbf{1})}{\mu_{\varphi} - \varphi} \) along with \( w' \bar{\mu} = \mu_{\varphi} \), we may express these as follows:

\[ C = \frac{2 (\mu_{\varphi} - \varphi)}{T \sigma_{\phi}^2} A^{-1} \sum_{t=1}^{T} Z_{t-1} \left[ 1 - \frac{\mu_{\varphi} - \varphi}{\sigma_{\phi}^2} \left( \mu_t w - \mu_{\varphi} \right) \right] w' \]  
\[ D = -\frac{(\mu_{\varphi} - \varphi)^2}{\sigma_{\phi}^4} ww' = -\frac{S^2_{\phi}}{\sigma_{\phi}^2} ww' \]  

(A.145)

To see that the canonical matrices have the same functional form, for the squared Sharpe Ratio, in the case of a given set of fixed weights \( w \) and for weights that are estimated so as to maximize the squared Sharpe Ratio, we note that the estimated portfolio mean may be written as

\[ \hat{\mu}_p = \mu_p + \frac{1}{T} \sum_{t=1}^{T} w'(\hat{\mu}_t - \mu_t) \]  

(A.146)

where the portfolio mean is \( \mu_p = \frac{1}{T} \sum_{t=1}^{T} w \mu_t \).
The portfolio variance may be written as the expected conditional variance plus the variance of the conditional mean:

\[ \sigma_p^2 = w'Vw + \frac{1}{T} \sum_{t=1}^{T} (w'\mu_t)^2 - \left( \frac{1}{T} \sum_{t=1}^{T} w'\mu_t \right)^2 \]  \hfill (A.147)

and may be estimated as

\[ \hat{\sigma}_p^2 = w'\hat{V}w + \frac{1}{T} \sum_{t=1}^{T} (w'\hat{\mu}_t)^2 - \left( \frac{1}{T} \sum_{t=1}^{T} w'\hat{\mu}_t \right)^2 \]

\[ \quad \quad = \sigma_p^2 + \frac{2}{T} \sum_{t=1}^{T} (w'\mu_t) w'(\hat{\mu}_t - \mu_t) - \left( \frac{2}{T} \sum_{t=1}^{T} w'\hat{\mu}_t \right) \left( \frac{1}{T} \sum_{t=1}^{T} w'(\hat{\mu}_t - \mu_t) \right) \]

\[ \quad \quad + w'(\hat{V} - V)w + O_p(1/T) \]

\[ \quad \quad = \sigma_p^2 + \frac{2}{T} \sum_{t=1}^{T} (w'\mu_t - \mu_p) w'(\hat{\mu}_t - \mu_t) + w'(\hat{V} - V)w + O_p(1/T) \]  \hfill (A.148)

The estimated squared Sharpe Ratio, given the zero-beta rate \( \varphi \), may then be expanded as follows:

\[ \hat{S}_p^2 = \frac{(\hat{\mu}_p - \varphi)^2}{\hat{\sigma}_p^2} = \left[ \frac{(\mu_p - \varphi) + (\hat{\mu}_p - \mu_p)}{\sigma_p^2 + (\hat{\sigma}_p^2 - \sigma_p^2)} \right]^2 = \frac{(\mu_p - \varphi)^2 + 2(\mu_p - \varphi)(\hat{\mu}_p - \mu_p)}{\sigma_p^2 \left[ 1 + (\hat{\sigma}_p^2 - \sigma_p^2) / \sigma_p^2 \right]} + O_p(1/T) \]

\[ \quad \quad = \frac{(\mu_p - \varphi)^2 + 2(\mu_p - \varphi)(\hat{\mu}_p - \mu_p) \left[ 1 - (\hat{\sigma}_p^2 - \sigma_p^2) / \sigma_p^2 \right]}{\sigma_p^2} + O_p(1/T) \]  \hfill (A.149)

\[ \quad \quad = \frac{(\mu_p - \varphi)^2}{\sigma_p^2} + \frac{2(\mu_p - \varphi)(\hat{\mu}_p - \mu_p)}{\sigma_p^2} - \frac{(\mu_p - \varphi)^2}{\sigma_p^4} (\hat{\sigma}_p^2 - \sigma_p^2) + O_p(1/T) \]

Substituting from \( (A.113) \) and \( (A.115) \) into \( (A.116) \) we find
\[
\hat{S}_p^2 = \frac{(\mu_p - \varphi)^2}{\sigma_p^2} + 2\left(\frac{\mu_p - \varphi}{\sigma_p^2}\right)(\hat{\mu}_p - \mu_p) - \frac{(\mu_p - \varphi)^2}{\sigma_p^4}(\hat{\sigma}_p^2 - \sigma_p^2) + \mathcal{O}_p(1/T)
\]
\[
= S_p^2 + \frac{2(\mu_p - \varphi)}{T\sigma_p^2} \sum_{t=1}^T w'(\hat{\mu}_t - \mu_t) - \frac{2(\mu_p - \varphi)^2}{T\sigma_p^4} \sum_{t=1}^T (w'_t \mu_t - \mu_p)w'(\hat{\mu}_t - \mu_t)
\]
\[
- \frac{(\mu_p - \varphi)^2}{\sigma_p^4} w'(\hat{\mu} - V)w + \mathcal{O}_p(1/T) \tag{A.150}
\]
\[
= S_p^2 + \frac{2(\mu_p - \varphi)}{T\sigma_p^2} \sum_{t=1}^T \left[1 - \frac{\mu_p - \varphi}{\sigma_p^2}(\mu'_t w - \mu_p)\right]w'(\hat{\mu}_t - \mu_t)
\]
\[
- \frac{(\mu_p - \varphi)^2}{\sigma_p^4} w'(\hat{\mu} - V)w + \mathcal{O}_p(1/T)
\]

for which we find estimates of the canonical matrices

\[
C = \frac{2(\mu_p - \varphi)A^{-1}}{T\sigma_p^2} \sum_{t=1}^T Z_{t-1} \left[1 - \frac{\mu_p - \varphi}{\sigma_p^2}(\mu'_t w - \mu_p)\right]w' \tag{A.151}
\]
\[
D = -\frac{(\mu_p - \varphi)^2}{\sigma_p^4} w w'
\]

which have the same functional form as in the case when the weights are estimated to maximize the squared Sharpe Ratio, completing the proof. \(\square\)

Proof of Corollary II:

To obtain expansions for \(\hat{\alpha}_1\), \(\hat{\alpha}_2\), and \(\hat{\alpha}_3\), we begin with \(\hat{\Lambda}_r\) to find

\[
\hat{\Lambda}_r = (\hat{\mu}_r + \hat{\mu}' + \hat{\mu})^{-1} = \left[\mu_r + \mu_r (\hat{\mu}_r - \mu_r)' + (\hat{\mu}_r - \mu_r)\mu_r' + (\hat{\mu} - V)^{-1} + \mathcal{O}_p(1/T)\right]
\]
\[
= \Lambda_r \left[I + \left[\mu_r (\hat{\mu}_r - \mu_r)' + (\hat{\mu}_r - \mu_r)\mu_r' + (\hat{\mu} - V)^{-1}\right] \Lambda_r \right]^{-1} + \mathcal{O}_p(1/T) \tag{A.152}
\]
\[
= \Lambda_r \left[I - \left[\mu_r (\hat{\mu}_r - \mu_r)' + (\hat{\mu}_r - \mu_r)\mu_r' + (\hat{\mu} - V)^{-1}\right] \Lambda_r \right] + \mathcal{O}_p(1/T)
\]
\[
= \Lambda_r - \Lambda_r \mu_r (\hat{\mu}_r - \mu_r)' \Lambda_r \Lambda_r (\hat{\mu}_r - \mu_r) \mu_r' \Lambda_r - \Lambda_r (\hat{\mu} - V) \Lambda_r + \mathcal{O}_p(1/T)
\]
and continue, expanding

\[ 1' \hat{\Lambda}_1 = 1' \Lambda_1 - 1' \Lambda_{\mu_t} (\hat{\mu}_t - \mu_t) \Lambda_1 - 1' \Lambda_{\mu} (\hat{\mu}_t - \mu_t) \mu' \Lambda_1 - 1' \Lambda_{\mu_t} (\hat{V} - V) \Lambda_1 + O_p (1/T) \]  

(A.153)

and

\[ \frac{1}{1' \hat{\Lambda}_1} = \frac{1}{1' \Lambda_1} + \frac{21' \Lambda_{\mu_t} (\hat{\mu}_t - \mu_t) + 1' \Lambda_{\mu} (\hat{V} - V) \Lambda_1}{(1' \Lambda_1)^2} + O_p (1/T) \]  

(A.154)

Next, we expand \( \hat{\alpha}_1 \) as follows

\[ \hat{\alpha}_1 = \frac{1}{T} \sum_{t=1}^{T} \frac{1}{1' \hat{\Lambda}_1} \]

\[ = \alpha_1 + \frac{1}{T} \sum_{t=1}^{T} \frac{21' \Lambda_{\mu_t} (\hat{\mu}_t - \mu_t)}{(1' \Lambda_1)^2} + \frac{1}{T} \sum_{t=1}^{T} \frac{1' \Lambda_{\mu} (\hat{V} - V) \Lambda_1}{(1' \Lambda_1)^2} + O_p (1/T) \]  

(A.155)

To expand \( \hat{\alpha}_2 \), we will need

\[ 1' \hat{\Lambda}_2 \hat{\mu}_t = 1' \Lambda_{\mu_t} + 1' \Lambda_{(\hat{\Lambda}_t - \Lambda_t)} \mu_t + 1' \Lambda_{(\hat{\mu}_t - \mu_t)} + O_p (1/T) \]

\[ = 1' \Lambda_{\mu_t} + 1' \Lambda_{(\hat{\Lambda}_t - \Lambda_t)} \mu_t + 1' \Lambda_{(\hat{\mu}_t - \mu_t)} \mu' \Lambda_t - \Lambda_t (\hat{V} - V) \Lambda_t \mu_t \]

(A.156)

\[ + 1' \Lambda_{(\hat{\mu}_t - \mu_t)} + O_p (1/T) \]

\[ = 1' \Lambda_{\mu_t} + \left(1' - 1' \Lambda_{\mu_t} \mu' \mu'_t \right) \Lambda_t (\hat{\mu}_t - \mu_t) - 1' \Lambda_{(\hat{V} - V)} \Lambda_t \mu_t + O_p (1/T) \]

and
\[
\begin{align*}
\frac{1'}{1} & = \frac{1'}{1}
+ \frac{1}{1'} \left[ \left( 1' - 1' \mu, \mu' - \mu', \mu, \mu, 1' \right) \Lambda, \mu, \mu, 1' \right] \\
& + \left( 1' \mu, \mu, 1' \mu, \mu, 1' \right) \left( 1' \mu, \mu, 1' \mu, \mu, 1' \right) + O_p \left( 1/T \right) \\
& = \frac{1'}{1'} \left( 1' - 1' \mu, \mu', \mu', \mu, 1' \right) + 2 \left( 1' \mu, \mu \right)^2 1' \Lambda, \mu, 1 + O_p \left( 1/T \right)
\end{align*}
\] (A.157)

from which we find

\[
\hat{\alpha}_2 = \frac{1}{T} \sum_{t=1}^T \frac{1'}{1'} \frac{\hat{\mu}_t}{1'}
 = \alpha_2 + \frac{1}{T} \sum_{t=1}^T \left( 1' \mu T \right) \left( 1' - 1' \mu, \mu', \mu', \mu, 1' \right) + 2 \left( 1' \mu, \mu \right)^2 1' \Lambda, \mu, 1 + O_p \left( 1/T \right)
\] (A.158)

To expand \( \hat{\alpha}_3 \), we will also need

\[
\hat{\mu}, \hat{\mu} = \mu' \Lambda, \mu + 2 \mu' \Lambda, (\hat{\mu} - \mu_1) + \mu' \left( \hat{\mu} - \Lambda, \mu, \mu, 1' \right) + O_p \left( 1/T \right)
\]

\[
= \mu' \Lambda, \mu + 2 \mu' \Lambda, (\hat{\mu} - \mu, \mu, \mu, 1') + \mu' \left[ -\Lambda, \mu, (\hat{\mu} - \mu, \mu, \mu, \mu, \mu, \mu, \mu, \mu, 1') \right] + O_p \left( 1/T \right)
\] (A.159)

as well as
\[
\begin{align*}
(\mathbf{1}'\hat{\Lambda},\hat{\mu},)^2 &= (\mathbf{1}'\Lambda,\mu,)^2 + 2(\mathbf{1}'\Lambda,\mu,)\mathbf{1}'(\hat{\Lambda}, - \Lambda, )\mu, + 2(\mathbf{1}'\Lambda,\mu,)\mathbf{1}'\Lambda,(\hat{\mu}, - \mu,) + O_p (1/T) \\
&= (\mathbf{1}'\Lambda,\mu,)^2 + 2(\mathbf{1}'\Lambda,\mu,)\mathbf{1}'\Lambda,(\hat{\mu}, - \mu,)\Lambda, - \Lambda, (\hat{\mu}, - \mu,)\mu,\Lambda, - \Lambda, (\hat{\Lambda}, - V )\Lambda,\mu, \\
&\quad + 2(\mathbf{1}'\Lambda,\mu,)\mathbf{1}'\Lambda,(\hat{\mu}, - \mu,) + O_p (1/T) \\
&= (\mathbf{1}'\Lambda,\mu,)^2 - 2(\mathbf{1}'\Lambda,\mu,)\mathbf{1}'\Lambda,(\hat{\mu}, - \mu,)\Lambda,\mu, - 2(\mathbf{1}'\Lambda,\mu,)\mathbf{1}'\Lambda,(\hat{\mu}, - \mu,)\mu,\Lambda, \mu, \\
&\quad - 2(\mathbf{1}'\Lambda,\mu,)\mathbf{1}'\Lambda,(\hat{\Lambda}, - V )\Lambda,\mu, + 2(\mathbf{1}'\Lambda,\mu,)\mathbf{1}'\Lambda,(\hat{\mu}, - \mu,) + O_p (1/T) \\
&= (\mathbf{1}'\Lambda,\mu,)^2 - 2(\mathbf{1}'\Lambda,\mu,)\mathbf{1}'\Lambda,(\hat{\mu}, - \mu,)\Lambda,\mu, - 2(\mathbf{1}'\Lambda,\mu,)\mu,\Lambda, \mu, + 2(\mathbf{1}'\Lambda,\mu,)\mathbf{1}'\Lambda,(\hat{\mu}, - \mu,) + O_p (1/T) \\
&\quad + 2(\mathbf{1}'\Lambda,\mu,)\mathbf{1}'\Lambda,(\hat{\mu}, - \mu,) - 2(\mathbf{1}'\Lambda,\mu,)\mathbf{1}'\Lambda,(\hat{\Lambda}, - V )\Lambda,\mu, + O_p (1/T) \\
&= (\mathbf{1}'\Lambda,\mu,)^2 + [-2(\mathbf{1}'\Lambda,\mu,)\mathbf{1}'\Lambda,(\hat{\mu}, - \mu,)\Lambda,\mu, - 2(\mathbf{1}'\Lambda,\mu,)\mu,\Lambda, \mu, + 2(\mathbf{1}'\Lambda,\mu,)\mathbf{1}'\Lambda,(\hat{\mu}, - \mu,) - 2(\mathbf{1}'\Lambda,\mu,)\mathbf{1}'\Lambda,(\hat{\Lambda}, - V )\Lambda,\mu, + O_p (1/T)
\end{align*}
\]

and

\[
(A.160)
\]
\[
\left( \Gamma \Lambda, \hat{\mu} \right)^2_{\hat{\Gamma} \Lambda, \hat{1}} = \left( \Gamma \Lambda, \mu \right)^2_{\Gamma \Lambda, \hat{1}} + \left( \Gamma \Lambda, \mu \right)^2_{\Gamma \Lambda, \hat{1}}\left( \begin{array}{c} 2(\Gamma \Lambda, \mu) (\Gamma \Lambda, \mu) (\hat{\mu} - \mu, \mu) + (\Gamma \Lambda, \hat{\mu} - \hat{\mu}, \hat{\mu}) \end{array} \right) + (\Gamma \Lambda, \hat{\mu} - \hat{\mu}, \hat{\mu}) \left( \hat{V} - \hat{V} \right) \Lambda, \hat{1} \right) \\
+ \frac{1}{\Gamma \Lambda, \hat{1}} \left[ 2(\Gamma \Lambda, \mu) (\Gamma \Lambda, \mu) (\mu, \mu, \mu) (\mu, \mu) \Lambda, (\hat{\mu}, \mu, \mu) - 2(\Gamma \Lambda, \mu) (\Gamma \Lambda, \hat{\mu} - \hat{\mu}, \hat{\mu}) \left( \hat{V} - \hat{V} \right) \Lambda, \mu \right] + O_p (1/T) \\
= \left( \Gamma \Lambda, \mu \right)^2_{\Gamma \Lambda, \hat{1}} + \frac{2(\Gamma \Lambda, \mu) (\Gamma \Lambda, \mu) (\mu, \mu, \mu) (\mu, \mu) \Lambda, (\hat{\mu}, \mu, \mu)\left( \hat{V} - \hat{V} \right) \Lambda, \mu}{(\Gamma \Lambda, \hat{1})^2} + 2(\Gamma \Lambda, \mu) (\Gamma \Lambda, \mu) (\mu, \mu, \mu) (\mu, \mu) \Lambda, (\hat{\mu}, \mu, \mu) \left( \hat{V} - \hat{V} \right) \Lambda, \mu \\
+ \frac{2(\Gamma \Lambda, \mu) (\Gamma \Lambda, \mu) (\mu, \mu, \mu) (\mu, \mu) \Lambda, (\hat{\mu}, \mu, \mu)\left( \hat{V} - \hat{V} \right) \Lambda, \mu}{(\Gamma \Lambda, \hat{1})^2} + O_p (1/T) \\
= \left( \Gamma \Lambda, \mu \right)^2_{\Gamma \Lambda, \hat{1}} + \frac{2(\Gamma \Lambda, \mu) (\Gamma \Lambda, \mu) (\mu, \mu, \mu) (\mu, \mu) \Lambda, (\hat{\mu}, \mu, \mu)\left( \hat{V} - \hat{V} \right) \Lambda, \mu}{(\Gamma \Lambda, \hat{1})^2} + 2(\Gamma \Lambda, \mu) (\Gamma \Lambda, \mu) (\mu, \mu, \mu) (\mu, \mu) \Lambda, (\hat{\mu}, \mu, \mu) \left( \hat{V} - \hat{V} \right) \Lambda, \mu \\
+ \frac{2(\Gamma \Lambda, \mu) (\Gamma \Lambda, \mu) (\mu, \mu, \mu) (\mu, \mu) \Lambda, (\hat{\mu}, \mu, \mu)\left( \hat{V} - \hat{V} \right) \Lambda, \mu}{(\Gamma \Lambda, \hat{1})^2} + O_p (1/T) \\
= \left( \Gamma \Lambda, \mu \right)^2_{\Gamma \Lambda, \hat{1}} + \frac{2(\Gamma \Lambda, \mu) (\Gamma \Lambda, \mu) (\mu, \mu, \mu) (\mu, \mu) \Lambda, (\hat{\mu}, \mu, \mu)\left( \hat{V} - \hat{V} \right) \Lambda, \mu}{(\Gamma \Lambda, \hat{1})^2} + 2(\Gamma \Lambda, \mu) (\Gamma \Lambda, \mu) (\mu, \mu, \mu) (\mu, \mu) \Lambda, (\hat{\mu}, \mu, \mu) \left( \hat{V} - \hat{V} \right) \Lambda, \mu \\
+ \frac{2(\Gamma \Lambda, \mu) (\Gamma \Lambda, \mu) (\mu, \mu, \mu) (\mu, \mu) \Lambda, (\hat{\mu}, \mu, \mu)\left( \hat{V} - \hat{V} \right) \Lambda, \mu}{(\Gamma \Lambda, \hat{1})^2} + O_p (1/T) \\
\tag{A.161}
\]

To find

\[
\hat{\alpha}_j = \frac{1}{T} \sum_{t=1}^T \hat{\mu}_t \left( \hat{\Lambda} - \hat{\Lambda} \hat{\mu} \hat{\Lambda} \right)_{\hat{\Gamma} \Lambda, \hat{1}} = \frac{1}{T} \sum_{t=1}^T \left( \hat{\mu}_t \hat{\Lambda}, \hat{\mu}_t - \hat{\mu}_t \hat{\Lambda} \hat{\mu} \hat{\Lambda} \right)_{\hat{\Gamma} \Lambda, \hat{1}} \\
= \alpha_j + \frac{2}{T} \sum_{t=1}^T \left( 1 - \mu, \mu, \mu \right) \mu, \Lambda, (\hat{\mu}, \mu, \mu) - \frac{1}{T} \sum_{t=1}^T \mu, \Lambda, (\hat{\mu}, \mu, \mu) \left( \hat{V} - \hat{V} \right) \Lambda, \mu, \\
- \frac{2}{T} \sum_{t=1}^T (\Gamma \Lambda, \mu) (\Gamma \Lambda, \mu) (\mu, \mu, \mu) (\mu, \mu) \Lambda, (\hat{\mu}, \mu, \mu) \left( \hat{V} - \hat{V} \right) \Lambda, \mu + O_p (1/T) \\
= \alpha_j + \frac{2}{T} \sum_{t=1}^T \left( 1 - \mu, \mu, \mu \right) \mu, (\Gamma \Lambda, \mu) (\Gamma \Lambda, \mu) (\mu, \mu, \mu) (\mu, \mu) \Lambda, (\hat{\mu}, \mu, \mu) \left( \hat{V} - \hat{V} \right) \Lambda, \mu \\
+ \frac{2}{T} \sum_{t=1}^T (\Gamma \Lambda, \mu) (\Gamma \Lambda, \mu) (\mu, \mu, \mu) (\mu, \mu) \Lambda, (\hat{\mu}, \mu, \mu) \left( \hat{V} - \hat{V} \right) \Lambda, \mu + O_p (1/T) \\
\tag{A.162}
\]
Next, we use commutativity of matrix multiplication within the trace operator to write
\[
\hat{\alpha}_3 = \alpha_3 + \frac{2}{T} \sum_{t=1}^{T} \left( (1 - \mu'_{t} \Lambda, \mu_{t}) \mu_{t} - (\mathbf{I}' \Lambda, \mu_{t}) \frac{\mathbf{I}' - (\mu'_{t} \Lambda, \mu_{t}) \mu_{t}'}{\mathbf{I}' \Lambda, \mathbf{1}} - \frac{(\mathbf{I}' \Lambda, \mu_{t})^3 \mathbf{I}'}{(\mathbf{I}' \Lambda, \mathbf{1})^2} \right) \Lambda_{t}(\hat{\mu}_{t} - \mu_{t}) + O_p(1/T) \tag{A.163}
\]
\[
- \text{tr} \left[ \frac{1}{T} \sum_{t=1}^{T} \left( \Lambda_{t} \mu_{t} \mu'_{t} + (\mathbf{I}' \Lambda, \mu_{t}) \frac{\mathbf{I}' - 2(\mathbf{I}' \Lambda, \mu_{t}) \mu_{t}'}{(\mathbf{I}' \Lambda, \mathbf{1})^2} \right) \Lambda_{t}(\hat{V} - V) \right] + O_p(1/T)
\]
To summarize the expansions of the portfolio coefficients:
\[
\hat{\alpha}_1 = \alpha_1 + \frac{1}{T} \sum_{t=1}^{T} \frac{2(\mathbf{I}' \Lambda, \mu_{t}) (\hat{\mu}_{t} - \mu_{t})}{(\mathbf{I}' \Lambda, \mathbf{1})^2} + \frac{1}{T} \sum_{t=1}^{T} \frac{\mathbf{I}' \Lambda_{t} (\hat{V} - V) \Lambda_{t} \mathbf{1}}{(\mathbf{I}' \Lambda, \mathbf{1})^2} + O_p(1/T) \tag{A.164}
\]
\[
= \alpha_1 + \sum_{t=1}^{T} C_{\alpha_1,t} (\hat{\mu}_{t} - \mu_{t}) + \text{tr} \left[ D_{\alpha_1} (\hat{V} - V) \right] + O_p(1/T)
\]
\[
\hat{\alpha}_2 = \alpha_2 + \frac{1}{T} \sum_{t=1}^{T} \left( \frac{\mathbf{I}' - (\mu'_{t} \Lambda, \mu_{t}) \mu_{t}'}{\mathbf{I}' \Lambda, \mathbf{1}} + \frac{2(\mathbf{I}' \Lambda, \mu_{t})^3 \mathbf{I}'}{(\mathbf{I}' \Lambda, \mathbf{1})^2} \right) \Lambda_{t}(\hat{\mu}_{t} - \mu_{t})
\]
\[
+ \frac{1}{T} \sum_{t=1}^{T} \frac{(\mathbf{I}' \Lambda, \mu_{t}) \mathbf{I}' - (\mathbf{I}' \Lambda, \mathbf{1}) \mu_{t}'}{(\mathbf{I}' \Lambda, \mathbf{1})^2} \Lambda_{t}(\hat{V} - V) \Lambda_{t} \mathbf{1} + O_p(1/T) \tag{A.165}
\]
\[
= \alpha_2 + \sum_{t=1}^{T} C_{\alpha_2,t} (\hat{\mu}_{t} - \mu_{t}) + \text{tr} \left[ D_{\alpha_2} (\hat{V} - V) \right] + O_p(1/T)
\]
\[
\hat{\alpha}_3 = \alpha_3 + \frac{2}{T} \sum_{t=1}^{T} \left( (1 - \mu'_{t} \Lambda, \mu_{t}) \mu_{t} - (\mathbf{I}' \Lambda, \mu_{t}) \frac{\mathbf{I}' - (\mu'_{t} \Lambda, \mu_{t}) \mu_{t}'}{\mathbf{I}' \Lambda, \mathbf{1}} - \frac{(\mathbf{I}' \Lambda, \mu_{t})^3 \mathbf{I}'}{(\mathbf{I}' \Lambda, \mathbf{1})^2} \right) \Lambda_{t}(\hat{\mu}_{t} - \mu_{t})
\]
\[
- \text{tr} \left[ \frac{1}{T} \sum_{t=1}^{T} \left( \Lambda_{t} \mu_{t} \mu'_{t} + (\mathbf{I}' \Lambda, \mu_{t}) \frac{\mathbf{I}' - 2(\mathbf{I}' \Lambda, \mu_{t}) \mu_{t}'}{(\mathbf{I}' \Lambda, \mathbf{1})^2} \right) \Lambda_{t}(\hat{V} - V) \right] + O_p(1/T) \tag{A.166}
\]
\[
= \alpha_3 + \sum_{t=1}^{T} C_{\alpha_3,t} (\hat{\mu}_{t} - \mu_{t}) + \text{tr} \left[ D_{\alpha_3} (\hat{V} - V) \right] + O_p(1/T)
\]
The estimates of the terms for the Theorem I are given by
\[
\mu^F = \frac{1}{T} \sum_{t=1}^{T} \mu_{t}^F \tag{A.167}
\]
\[
\gamma_1 = \frac{1}{T} \sum_{t=1}^{T} \frac{1}{\mathbf{I}' \Lambda, \mathbf{1}} \tag{A.168}
\]
\[ \gamma_{\mu} = \frac{1}{T} \sum_{t=1}^{T} \frac{1' \Lambda_{t} \mu_{t}}{1' \Lambda_{t} \lambda_{t}} \]  
(A.169)

\[ \gamma_{F} = \frac{1}{T} \sum_{t=1}^{T} \frac{1' \Lambda_{t} \mathbb{E}(R_{F_{t}} \mid Z_{t-1})}{1' \Lambda_{t} \lambda_{t}} = \frac{1}{T} \sum_{t=1}^{T} \frac{1' \Lambda_{t} V_{F_{t}} + 1' \Lambda_{t} \mu_{t} \mu_{t}'}{1' \Lambda_{t} \lambda_{t}} \]  
(A.170)

\[ \gamma_{\mu'} = \frac{1}{T} \sum_{t=1}^{T} \mu_{t}' \Omega_{t} \mu_{t} \]  
(A.171)

\[ \gamma_{\mu F} = \frac{1}{T} \sum_{t=1}^{T} \mu_{t}' \Omega_{t} \mathbb{E}(R_{F_{t}} \mid Z_{t-1}) = \frac{1}{T} \sum_{t=1}^{T} \left( \mu_{t}' \Omega_{t} V_{F_{t}} + \mu_{t}' \Omega_{t} \mu_{t} \left( \mu_{t}' \right)^{2} \right) \]  
(A.172)

\[ \gamma_{FF} = \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}(R_{F_{t}} \mid Z_{t-1}) \Omega_{t} \mathbb{E}(R_{F_{t}} \mid Z_{t-1}) \]  
(A.173)

\[ = \frac{1}{T} \sum_{t=1}^{T} \left[ V_{F_{t}}' \Omega_{t} V_{F_{t}} + 2 \mu_{t}' \Omega_{t} V_{F_{t}} \mu_{t}' + \mu_{t}' \Omega_{t} \mu_{t} \left( \mu_{t}' \right)^{2} \right] \]

\[ C_{\alpha_{1}, t} = \frac{2(1' \Lambda_{t} \mu_{t}) 1' \Lambda_{t}}{T (1' \Lambda_{t} \lambda_{t})^{2}} \]  
(A.174)

\[ C_{\alpha_{1}} = \frac{2A^{-1}}{T} \sum_{t=1}^{T} Z_{t-1} \frac{(1' \Lambda_{t} \mu_{t}) 1' \Lambda_{t}}{(1' \Lambda_{t} \lambda_{t})^{2}} \]

\[ D_{\alpha_{1}} = \frac{1}{T} \sum_{t=1}^{T} \frac{\Lambda_{t} 1' \Lambda_{t} \lambda_{t}}{(1' \Lambda_{t} \lambda_{t})^{2}} \]  
(A.175)

\[ C_{\alpha_{2}, t} = \frac{1}{T} \left( \frac{1' - (\mu_{t}' \Lambda_{t} \mu_{t}) 1' - (1' \Lambda_{t} \mu_{t}) \mu_{t}'}{1' \Lambda_{t} \lambda_{t}} + \frac{2(1' \Lambda_{t} \mu_{t})^{2} 1'}{(1' \Lambda_{t} \lambda_{t})^{2}} \right) \Lambda_{t} \]

\[ C_{\alpha_{2}} = \frac{A^{-1}}{T} \sum_{t=1}^{T} Z_{t-1} \left( \frac{1' - (\mu_{t}' \Lambda_{t} \mu_{t}) 1' - (1' \Lambda_{t} \mu_{t}) \mu_{t}'}{1' \Lambda_{t} \lambda_{t}} + \frac{2(1' \Lambda_{t} \mu_{t})^{2} 1'}{(1' \Lambda_{t} \lambda_{t})^{2}} \right) \Lambda_{t} \]

\[ D_{\alpha_{2}} = \frac{1}{T} \sum_{t=1}^{T} \Lambda_{t} \left( \frac{1' \Lambda_{t} \mu_{t}}{1' \Lambda_{t} \lambda_{t}} \right) \left( 1' - (1' \Lambda_{t} \lambda_{t}) 1' \mu_{t} \right) \Lambda_{t} \]  
(A.176)

\[ C_{\alpha_{3}, t} = 2 \left( \frac{(1 - \mu_{t}' \Lambda_{t} \mu_{t}) \mu'_{t} - (1' \Lambda_{t} \mu_{t}) 1' - (1' \Lambda_{t} \mu_{t}) \mu'_{t}}{1' \Lambda_{t} \lambda_{t}} - \frac{(1' \Lambda_{t} \mu_{t})^{3} 1'}{(1' \Lambda_{t} \lambda_{t})^{2}} \right) \Lambda_{t} \]

\[ C_{\alpha_{3}} = \frac{2A^{-1}}{T} \sum_{t=1}^{T} Z_{t-1} \left( \frac{(1 - \mu_{t}' \Lambda_{t} \mu_{t}) \mu'_{t} - (1' \Lambda_{t} \mu_{t}) 1' - (1' \Lambda_{t} \mu_{t}) \mu'_{t}}{1' \Lambda_{t} \lambda_{t}} - \frac{(1' \Lambda_{t} \mu_{t})^{3} 1'}{(1' \Lambda_{t} \lambda_{t})^{2}} \right) \Lambda_{t} \]

\[ D_{\alpha_{3}} = -\frac{1}{T} \sum_{t=1}^{T} \Lambda_{t} \left( \mu_{t} \mu'_{t} + (1' \Lambda_{t} \mu_{t}) \frac{(1' \Lambda_{t} \mu_{t}) 1' - 2(1' \Lambda_{t} \lambda_{t}) 1' \mu'_{t}}{(1' \Lambda_{t} \lambda_{t})^{2}} \right) \Lambda_{t} \]
The estimated squared Sharpe Ratio is

$$\hat{S}_\varphi^2 = \frac{\hat{\alpha}_2^2 + \hat{\alpha}_4 \hat{\alpha}_3 - 2 \varphi \hat{\alpha}_2 + \varphi^2 (1 - \hat{\alpha}_3)}{\hat{\alpha}_1 (1 - \hat{\alpha}_3) - \hat{\alpha}_2^2}. \tag{A.177}$$

This estimate may be found by plugging in consistent estimates for the $\alpha$’s or by computing the estimated portfolio weights as a function of $Z$, applying these to the asset returns and computing the unconditional mean and variance of the resulting portfolio. In simulations not reported here, we find that either approach produces similar results. Equation (A.43) follows from maximizing

$$\hat{S}_\varphi^2 = \left( \hat{\mu}_p - \varphi \right)^2 = \left( \hat{\mu}_p - \varphi \right)^2 \frac{2 \hat{\alpha}_2}{\hat{\alpha}_3} \hat{\mu}_p + \frac{1 - \hat{\alpha}_3}{\hat{\alpha}_3} \hat{\mu}_p^2 \tag{A.178}$$

with respect to the mean $\hat{\mu}_p$. We expand (A.43) to obtain
\[
\hat{S}_\theta^2 = S_\theta^2 + \frac{\alpha_1 (1 - \alpha_3) - \alpha_2^*}{\alpha_1 (1 - \alpha_3) - \alpha_2^*} \frac{2 \alpha_2 (\hat{\alpha}_2 - \alpha_3) + \alpha_1 (\hat{\alpha}_1 - \alpha_3) + (\hat{\alpha}_1 - \alpha_3) \alpha_3 - 2 \varphi (\hat{\alpha}_2 - \alpha_3) - \varphi^2 (\hat{\alpha}_3 - \alpha_3)}{\left[ \alpha_1 (1 - \alpha_3) - \alpha_2^* \right]^2} \\
- \frac{\alpha_2^* + \alpha_1 \alpha_3 - 2 \varphi \alpha_2 + \varphi^2 (1 - \alpha_3)}{\left[ \alpha_1 (1 - \alpha_3) - \alpha_2^* \right]^2} \left[ -\alpha_1 (\hat{\alpha}_2 - \alpha_3) + (\hat{\alpha}_1 - \alpha_3) (1 - \alpha_3) - 2 \alpha_2 (\hat{\alpha}_2 - \alpha_2) \right] + O_p(1/T) \\
= S_\theta^2 + \frac{\alpha_1 (1 - \alpha_3) - \alpha_2^*}{\alpha_1 (1 - \alpha_3) - \alpha_2^*} \frac{\alpha_1 (\hat{\alpha}_1 - \alpha_1)}{\left[ \alpha_1 (1 - \alpha_3) - \alpha_2^* \right]^2} \\
+ 2 \frac{\alpha_1 (1 - \alpha_3) - \alpha_2^*}{\alpha_1 (1 - \alpha_3) - \alpha_2^*} \frac{\alpha_2 (\hat{\alpha}_2 - \alpha_3)}{\left[ \alpha_1 (1 - \alpha_3) - \alpha_2^* \right]^2} \\
+ \frac{\alpha_1 (1 - \alpha_3) - \alpha_2^*}{\alpha_1 (1 - \alpha_3) - \alpha_2^*} \frac{\alpha_3 (\hat{\alpha}_3 - \alpha_3)}{\left[ \alpha_1 (1 - \alpha_3) - \alpha_2^* \right]^2} + O_p(1/T) \tag{A.179}
\]

and then gather terms to find

\[
\hat{S}_\theta^2 = S_\theta^2 + \frac{\alpha_1 (1 - \alpha_3) - \alpha_2^*}{\alpha_1 (1 - \alpha_3) - \alpha_2^*} \frac{\alpha_3}{\left[ \alpha_1 (1 - \alpha_3) - \alpha_2^* \right]^2} \left[ \alpha_2^* + \alpha_1 \alpha_3 - 2 \varphi \alpha_2 + \varphi^2 (1 - \alpha_3) \right] (1 - \alpha_3) (\hat{\alpha}_1 - \alpha_1) \\
+ 2 \frac{\alpha_1 (1 - \alpha_3) - \alpha_2^*}{\alpha_1 (1 - \alpha_3) - \alpha_2^*} \frac{\alpha_2 (\hat{\alpha}_2 - \alpha_3)}{\left[ \alpha_1 (1 - \alpha_3) - \alpha_2^* \right]^2} \\
+ \frac{\alpha_1 (1 - \alpha_3) - \alpha_2^*}{\alpha_1 (1 - \alpha_3) - \alpha_2^*} \frac{\alpha_3 (\hat{\alpha}_3 - \alpha_3)}{\left[ \alpha_1 (1 - \alpha_3) - \alpha_2^* \right]^2} + O_p(1/T) \tag{A.180}
\]

and simplify to obtain

\[
\hat{S}_\theta^2 = S_\theta^2 + \frac{(\alpha_2 - \varphi (1 - \alpha_3))^2}{\alpha_1 (1 - \alpha_3) - \alpha_2^*} (\hat{\alpha}_1 - \alpha_1) + \frac{2 (\alpha_1 - \varphi \alpha_2) (\alpha_2 - \varphi (1 - \alpha_3))^2}{\left[ \alpha_1 (1 - \alpha_3) - \alpha_2^* \right]^2} (\hat{\alpha}_2 - \alpha_2) \\
+ \frac{(\alpha_1 - \varphi \alpha_2)^2}{\left[ \alpha_1 (1 - \alpha_3) - \alpha_2^* \right]^2} (\hat{\alpha}_3 - \alpha_3) + O_p(1/T) \tag{A.181}
\]
so that the asymptotic variance of \( \hat{S}_q^2 \) will follow from the Theorem I together with the following choices for canonical matrices \( C \) and \( D \):

\[
C = -\left[ \alpha_2 - \varphi(1 - \alpha_3) \right]^2 C_{\alpha_2} + 2(\alpha_1 - \varphi \alpha_2) \left[ \alpha_2 - \varphi(1 - \alpha_3) \right] C_{\alpha_2} + (\alpha_1 - \varphi \alpha_2)^2 C_{\alpha_2}
\]

\[
D = -\left[ \alpha_2 - \varphi(1 - \alpha_3) \right]^2 D_{\alpha_2} + 2(\alpha_1 - \varphi \alpha_2) \left[ \alpha_2 - \varphi(1 - \alpha_3) \right] D_{\alpha_2} + (\alpha_1 - \varphi \alpha_2)^2 D_{\alpha_2}
\]

(A.182)

completing the proof. \( \square \)

**Proof of Corollary III:**

Our model for the joint distribution of the asset returns and the factor will nest the original specification, \( R_t = \mu_t + \varepsilon_t = \delta'Z_{t-1} + \varepsilon_t \) with constant conditional covariance matrix \( V \) of \( R_t \) given the lagged instruments \( Z_{t-1} \), within the extended specification \( R_t^* = \mu_t^* + \varepsilon_t^* = \delta''Z_{t-1} + \varepsilon_t^* \) by appending \( F \) to the list of asset returns. Specifically, we define the extended \( N \times 1 \) vectors \( R_t^* \equiv (R_t', F_t')' = (R_t^1, ..., R_t^n, F_t) \)\( \mu_t^* \equiv (\mu_t^1, ..., \mu_t^n, \mu_t^F) \)' and \( \varepsilon_t^* \equiv (\varepsilon_t^1, ..., \varepsilon_t^n, \varepsilon_t^F) \)' We define the extended \( k \times (n + 1) \) matrix

\[
\delta^* = \begin{bmatrix} \delta & \delta_F \end{bmatrix}
\]

(A.183)

where \( \delta \) is the same \( L \times N \) matrix of regression coefficients as before, and \( \delta_F \) contains the \( L \) regression coefficients for \( F \) given \( Z \). We also define the extended \( (N + 1) \times (N + 1) \) conditional covariance matrix

\[
V^* = \begin{bmatrix} V & V_F \\ V_F' & \sigma_{F|Z}^2 \end{bmatrix}
\]

(A.184)
where $V$ is the same $N \times N$ conditional covariance matrix as before, $V_F \equiv \text{Cov}(R_t F_t | Z_{t-1})$ is $N \times 1$, and $\sigma^2_{\epsilon_t F} \equiv \text{Var}(F_t | Z_{t-1})$ is a scalar. We assume that the unobserved $\epsilon_t^*$ are independent and identically distributed in this extended model, with mean zero and covariance matrix $V^*$. We note that

$$E(R_t F_t | Z_{t-1}) = \text{Cov}(R_t F_t | Z_{t-1}) + \mu_t \mu_t^F = V_F + \mu_1 \mu_1^F$$  \hspace{1cm} (A.185)$$

Using Ferson, Siegel, and Xu (2006) Equation (6) and Corollary to Proposition 2, the maximal correlation portfolio is

$$w_t' = \frac{1' \Lambda_t}{1' \Lambda_t 1} - \left[ \lambda_1 \mu_1' + \lambda_2 E(F_t R_t' | Z_{t-1}) \right] \left( \Lambda_t - \frac{\Lambda_t 1' \Lambda_t}{1' \Lambda_t 1} \right)$$  \hspace{1cm} (A.186)$$

where

$$\Lambda_t \equiv (\mu_1, \mu_1') + V$$  \hspace{1cm} (A.187)$$

$$\lambda_1 \equiv -\frac{\gamma_1 (\mu_2 - \gamma_2 \mu_F) + \gamma_2 \gamma_F}{\gamma_2 (\mu_2 - \gamma_2 \mu_F) + \gamma_2 (\gamma_2 \mu_2 - 1)}$$  \hspace{1cm} (A.188)$$

$$\lambda_2 \equiv -\frac{\gamma_1 (\gamma_2 \mu_2 - 1) - \gamma_2 \gamma^2}{\gamma_2 (\mu_2 - \gamma_2 \mu_F) + \gamma_2 (\gamma_2 \mu_2 - 1)}$$  \hspace{1cm} (A.189)$$

Consistent estimates of the model parameters are:

$$\mu_1^F \equiv \frac{1}{T} \sum_{t=1}^{T} \mu_1^F$$  \hspace{1cm} (A.190)$$

$$\gamma_1 \equiv \frac{1}{T} \sum_{t=1}^{T} \frac{1}{1' \Lambda_t 1}$$  \hspace{1cm} (A.191)$$

$$\gamma_2 \equiv \frac{1}{T} \sum_{t=1}^{T} \frac{1' \Lambda_t \mu_t}{1' \Lambda_t 1}$$  \hspace{1cm} (A.192)$$

$$\gamma_F \equiv \frac{1}{T} \sum_{t=1}^{T} \frac{1' \Lambda_t E(R_t F_t | Z_{t-1})}{1' \Lambda_t 1} = \frac{1}{T} \sum_{t=1}^{T} \frac{1' \Lambda_t V_F + 1' \Lambda_t \mu_1^F}{1' \Lambda_t 1}$$  \hspace{1cm} (A.193)$$
\[
\Omega_t \equiv \Lambda_t - \frac{\Lambda_t \Lambda'_{t} \Lambda_t}{1' \Lambda_t 1} \quad (A.194)
\]

\[
\gamma_{\mu \mu} \equiv \frac{1}{T} \sum_{t=1}^{T} \mu'_t \Omega_t \mu_t \quad (A.195)
\]

\[
\gamma_{\mu F} \equiv \frac{1}{T} \sum_{t=1}^{T} \mu'_t \Omega_F E\left(R_t F_t | Z_{t-1}\right) = \frac{1}{T} \sum_{t=1}^{T} \left(\mu'_t \Omega_F V_F + \mu'_t \Omega_F \mu_t F_t\right) \quad (A.196)
\]

\[
\gamma_{FF} \equiv \frac{1}{T} \sum_{t=1}^{T} E\left(R'_t F_t | Z_{t-1}\right) \Omega_F E\left(R'_t F_t | Z_{t-1}\right) = \frac{1}{T} \sum_{t=1}^{T} \left[V'_t \Omega_F V_F + 2 \mu'_t \Omega_F V_F \mu_t + \mu'_t \Omega_F \mu_t \left(\mu_t F_t\right)^2 \right] \quad (A.197)
\]

Canonical matrices for the squared Sharpe Ratio with respect to the zero-beta rate \( \varphi \), 
\[
S_p^2 = (\mu_p - \varphi)^2 / \sigma_p^2,
\]
will initially be developed separately for \( \mu_p \) and for \( \sigma_p^2 \), and then combined. For portfolio returns \( R_{p,t} = w'_t R_t \), the portfolio mean return is 
\[
\mu_p = \frac{1}{T} \sum_{t=1}^{T} w'_t \mu_t,
\]
and the conditional variance of \( R_{p,t} \) is 
\[
\sigma_p^2 \equiv \text{Var}\left(R_{p,t} | Z_{t-1}\right) = w'_t V w_t \]
using the conditional covariance matrix \( V \) of \( R_t \) given \( Z_{t-1} \). The variance \( \sigma_p^2 \) of the portfolio \( R_p \) may then be expressed as the expected conditional variance plus the variance of the conditional portfolio means \( w'_t \mu_t \), then developed as follows:

\[
\sigma_p^2 = \frac{1}{T} \sum_{t=1}^{T} \left[ \sigma_{p,t}^2 + (w'_t \mu_t)^2 \right] - \left( \frac{1}{T} \sum_{t=1}^{T} w'_t \mu_t \right)^2
\]

\[
= \frac{1}{T} \sum_{t=1}^{T} \left[ w'_t V w_t + (w'_t \mu_t)^2 \right] - \mu_p^2
\]

\[
= \frac{1}{T} \sum_{t=1}^{T} w'_t (V + \mu_t \mu'_t) w_t - \mu_p^2
\]

\[
= \frac{1}{T} \sum_{t=1}^{T} w'_t \Lambda_t^{-1} w_t - \mu_p^2
\]

For the weights as specified in (A.153) the portfolio mean is
\[ \mu_p = \frac{1}{T} \sum_{t=1}^{T} w_t \mu_t = \gamma - \lambda_1 \gamma \mu - \lambda_2 \gamma_{\mu F} \]  
\hspace{2cm} (A.199)

and the portfolio variance from (A.165) is

\[ \sigma_p^2 = \frac{1}{T} \sum_{t=1}^{T} w_t^\prime \Lambda_t^{-1} w_t - \mu_p^2 \]  
\hspace{2cm} (A.200)

which we will simplify using the fact that

\[ w_t^\prime \Lambda_t^{-1} w_t = \frac{1}{1^\prime \Lambda_t 1} \left[ \lambda_1 \mu_t^\prime + \lambda_2 E (F_t R_t^\prime | Z_{t-1}) \right] \left[ \Lambda_t - \frac{\Lambda_t 1 1^\prime \Lambda_t}{1^\prime \Lambda_t 1} \right] \left[ \lambda_1 \mu_t + \lambda_2 E (F_t R_t | Z_{t-1}) \right] \]  
\hspace{2cm} (A.201)

because

\[ 1^\prime \left( \Lambda_t - \frac{\Lambda_t 1 1^\prime \Lambda_t}{1^\prime \Lambda_t 1} \right) = 0 \]  
\hspace{2cm} (A.202)

and

\[ \left( \Lambda_t - \frac{\Lambda_t 1 1^\prime \Lambda_t}{1^\prime \Lambda_t 1} \right) \Lambda_t^{-1} \left( \Lambda_t - \frac{\Lambda_t 1 1^\prime \Lambda_t}{1^\prime \Lambda_t 1} \right) = \left( \Lambda_t - \frac{\Lambda_t 1 1^\prime \Lambda_t}{1^\prime \Lambda_t 1} \right) \]  
\hspace{2cm} (A.203)

to find that

\[ \sigma_p^2 = \gamma_1 + \lambda_1^2 \gamma \mu - 2 \lambda_1 \lambda_2 \gamma_{\mu F} + \lambda_2^2 \gamma_{FF} - \mu_p^2 \]  
\hspace{2cm} (A.204)

To facilitate the expansions that follow, we define the following scalars:

\[ \theta_t \equiv \frac{1}{1^\prime \Lambda_t 1} \]  
\hspace{2cm} (A.205)

\[ \theta_{\mu F} = 1^\prime \Lambda_t \mu_t \]  
\hspace{2cm} (A.206)

\[ \theta_{\mu F} \equiv \mu_t^\prime \Lambda_t \mu_t \]  
\hspace{2cm} (A.207)

\[ \theta_{F F} \equiv 1^\prime \Lambda_t V_{FF} \]  
\hspace{2cm} (A.208)
\[ \theta_{FF,t} = V'_t \Lambda_t V_F \]  
\[ \theta_{\mu F,t} = \mu'_t \Lambda_t V_F \]  
\[ \theta_{\mu \Omega_t,t} = \mu'_t \Omega_t \mu_t = \mu'_t \left( \Lambda_t - \frac{\Lambda_t \Omega_t \Lambda_t}{\Lambda_t \Omega_t \Lambda_t} \right) \mu_t = \theta_{\mu \Omega_t,t} - \theta_t \theta_{\mu,t} \]  
\[ \theta_{\mu \Omega F,t} = \mu'_t \Omega_t V_F = \mu'_t \left( \Lambda_t - \frac{\Lambda_t \Omega_t \Lambda_t}{\Lambda_t \Omega_t \Lambda_t} \right) V_F = \theta_{\mu \Omega F,t} - \theta_t \theta_{\mu \Omega,t} \]  
\[ \theta_{F \Omega F,t} = V'_t \Omega_t V_F = V'_t \left( \Lambda_t - \frac{\Lambda_t \Omega_t \Lambda_t}{\Lambda_t \Omega_t \Lambda_t} \right) V_F = \theta_{F \Omega F,t} - \theta_t \theta_{F,t} \]  

We may then write the consistent estimators as:

\[ \gamma_t = \frac{1}{T} \sum_{t=1}^{T} \theta_t \]  
\[ \gamma_\mu = \frac{1}{T} \sum_{t=1}^{T} \theta_t \theta_{\mu,t} \]  
\[ \gamma_F = \frac{1}{T} \sum_{t=1}^{T} \theta_t \left( \theta_{F,t} + \theta_{\mu,t} \mu_t^F \right) \]  
\[ \gamma_{\mu \mu} = \frac{1}{T} \sum_{t=1}^{T} \theta_{\mu \Omega_t,t} \]  
\[ \gamma_{\mu F} = \frac{1}{T} \sum_{t=1}^{T} \left( \mu'_t \Omega_t V_F + \mu'_t \Omega_t \mu_t^F \right) = \frac{1}{T} \sum_{t=1}^{T} \left( \theta_{\mu \Omega F,t} + \theta_{\mu \Omega,t} \mu_t^F \right) \]  
\[ \gamma_{FF} = \frac{1}{T} \sum_{t=1}^{T} \left[ V'_t \Omega_t V_F + 2 \mu'_t \Omega_t V_F \mu_t^F + \mu'_t \Omega_t \mu_t \left( \mu_t^F \right)^2 \right] \]  
\[ = \frac{1}{T} \sum_{t=1}^{T} \left[ \theta_{F \Omega F,t} + \theta_{\mu \Omega F,t} \mu_t^F + \theta_{\mu \Omega,t} \left( \mu_t^F \right)^2 \right] \]  

We begin by expanding

\[ \hat{\nu}_t = \frac{1}{\Lambda_t} = 0 + 2 \theta_t \theta_{\mu,t} \Lambda_t \left( \hat{\mu}_t - \mu_t \right) + Tr \left[ \theta_t \Lambda_t \Omega_t \left( \hat{V} - V \right) \right] + O_p \left( 1/T \right) \]
which follows from (A.27). We continue by using (A.25) to find

\[ \hat{\theta}_{\mu,j} = 1' \hat{\Lambda}_j \hat{\mu}_j = \theta_{\mu,j} + 1' \Lambda_j (\hat{\mu}_j - \mu_j) + 1' \left( \hat{\Lambda}_j - \Lambda_j \right) \mu_j + O_p (1/T) \]

\[ = \theta_{\mu,j} + 1' \Lambda_j (\hat{\mu}_j - \mu_j) + 1' \left[ -\Lambda_j \mu_j (\hat{\mu}_j - \mu_j)' \Lambda_j - \Lambda_j (\hat{\mu}_j - \mu_j) \mu_j' \Lambda_j - \Lambda_j \left( \hat{V} - V \right) \Lambda_j \right] \mu_j + O_p (1/T) \]

\[ = \theta_{\mu,j} + (1' - \theta_{\mu,j} \mu_j) \Lambda_j (\hat{\mu}_j - \mu_j) + tr \left[ -\Lambda_j \mu_j 1' \Lambda_j (\hat{V} - V) \right] + O_p (1/T) \quad (A.221) \]

\[ \hat{\theta}_{\mu,j} = \hat{\mu}_j' \hat{\Lambda}_j \hat{\mu}_j = \theta_{\mu,j} + 2 \mu_j' \Lambda_j (\hat{\mu}_j - \mu_j) + \mu_j' \left[ -\Lambda_j \mu_j (\hat{\mu}_j - \mu_j)' \Lambda_j - \Lambda_j (\hat{\mu}_j - \mu_j) \mu_j' \Lambda_j - \Lambda_j \left( \hat{V} - V \right) \Lambda_j \right] \mu_j + O_p (1/T) \]

\[ = \theta_{\mu,j} + (2 \mu_j' - 2 \theta_{\mu,j} \mu_j) \Lambda_j (\hat{\mu}_j - \mu_j) + tr \left[ -\Lambda_j \mu_j \mu_j' \Lambda_j (\hat{V} - V) \right] + O_p (1/T) \quad (A.222) \]

\[ \hat{\theta}_{\mu,j} = 1' \hat{\Lambda}_j \hat{V}_j = \theta_{\mu,j} + 1' (\hat{\Lambda}_j - \Lambda_j) V_j + 1' \Lambda_j \left( \hat{V}_j - V_j \right) + O_p (1/T) \]

\[ = \theta_{\mu,j} + 1' \left[ -\Lambda_j \mu_j (\hat{\mu}_j - \mu_j)' \Lambda_j - \Lambda_j (\hat{\mu}_j - \mu_j) \mu_j' \Lambda_j - \Lambda_j \left( \hat{V} - V \right) \Lambda_j \right] V_j \]

\[ + 1' \Lambda_j \left( \hat{V}_j - V_j \right) + O_p (1/T) \quad (A.223) \]

\[ = \theta_{\mu,j} + ( - \theta_{\mu,j} V_j' - \theta_{\mu,j} \mu_j) \Lambda_j (\hat{\mu}_j - \mu_j) + tr \left[ -\Lambda_j V_j \mu_j' \Lambda_j (\hat{V} - V) \right] \]

\[ + 1' \Lambda_j \left( \hat{V}_j - V_j \right) + O_p (1/T) \]

\[ \hat{\theta}_{\mu,j} = \hat{V}_j' \hat{\Lambda}_j \hat{V}_j = \theta_{\mu,j} + 2 V_j' \Lambda_j \left( \hat{V}_j - V_j \right) + V_j' \left( \hat{\Lambda}_j - \Lambda_j \right) V_j + O_p (1/T) \quad (A.224) \]

\[ = \theta_{\mu,j} + 2 V_j' \Lambda_j \left( \hat{V}_j - V_j \right) \]

\[ + V_j' \left[ -\Lambda_j \mu_j (\hat{\mu}_j - \mu_j)' \Lambda_j - \Lambda_j (\hat{\mu}_j - \mu_j) \mu_j' \Lambda_j - \Lambda_j \left( \hat{V} - V \right) \Lambda_j \right] V_j + O_p (1/T) \]

\[ = \theta_{\mu,j} - 2 \theta_{\mu,j} V_j' \Lambda_j (\hat{\mu}_j - \mu_j) + tr \left[ -\Lambda_j V_j V_j' \Lambda_j (\hat{V} - V) \right] \]

\[ + 2 V_j' \Lambda_j \left( \hat{V}_j - V_j \right) + O_p (1/T) \]
\[
\hat{\theta}_{F,t} = \hat{\mu}_t \hat{\Lambda}_i \hat{V}_F = \theta_{F,t} + \mu_t' (\hat{\Lambda}_t - \Lambda_t) V_F + V_t' \Lambda_t (\hat{\mu}_t - \mu_t) + \mu_t' \Lambda_t (\hat{V}_F - V_F) + O_p(1/T) \tag{A.225}
\]

To define the matrices to be used with the Theorem I applied to the extended specification, the following vector and matrix operators will be helpful. The first operator simply appends an additional scalar to a column vector, increasing its dimension from \( n \) to \( n + 1 \):

\[
\mathcal{L} \left[ (a_1, \ldots, a_n)' , b \right] = (a_1, \ldots, a_n, b)'
\tag{A.226}
\]

while the second operator enlarges a given \( n \times n \) matrix \( M \), along with a \( 1 \times n \) row vector \( X \), to size \((n + 1) \times (n + 1)\) by placing \( X \) below \( M \) and inserting a column of \( n + 1 \) zeros at the right, as follows:

\[
\mathcal{M} (M, X) \equiv \begin{bmatrix} M & 0 \\ X & 0 \end{bmatrix}
\tag{A.227}
\]

which will allow us to write, for example,

\[
\text{tr} \left[ \mathcal{M}(M, X)(\hat{V}^* - V^*) \right] = \text{tr} \left[ \begin{bmatrix} M & 0 \\ X & 0 \end{bmatrix} \begin{bmatrix} \hat{V} - V & \hat{V}_F - V_F \\ \hat{V}_F - V_F & \hat{V}_F - V_F \end{bmatrix} \right]
\]

\[
= \text{tr} \left[ \begin{bmatrix} M(\hat{V} - V) & M(\hat{V}_F - V_F) \\ X(\hat{V} - V) & X(\hat{V}_F - V_F) \end{bmatrix} \right] \tag{A.228}
\]

We are now ready to specify the matrices for the Theorem I, which we will define in sequence. We begin by defining the matrices for the expansion of \( \hat{\mu}^F = \mu^F + \frac{1}{T} \sum_{t=1}^T (\hat{\mu}_t^F - \mu_t^F) \), which has no \( \hat{V}^* - V^* \) term, as follows:
\[ C^F \equiv A^{-1} \sum_{t=1}^{T} Z_t \left( C_t^F \right)' \quad D^F \equiv \mathcal{M}(0,0) \]  
(A.229)

where \( C_t^F \equiv \mathcal{L}(0,1/T) \).

We next define \( C_{0t}' \) and \( D_{0t} \) (flattening the subscript levels for readability) for \( \theta_j \equiv \frac{1}{1' \Lambda_j 1} \), for which we append zeros (because there are no factor terms in \( \theta_j \)) to find

\[ C_{0t} \equiv \mathcal{L}\left( 20 \theta_{\mu,t} 1' \Lambda_j / T, 0 \right) \quad D_{0t} \equiv \mathcal{M}(\theta_j \Lambda_j, 0) \]  
(A.230)

Proceeding similarly, we find

\[ C_{0\mu} \equiv \mathcal{L}\left( (1' - \theta_{\mu,t} 1' - \theta_{\mu,t} \mu_t') \Lambda_j / T, 0 \right) \quad D_{0\mu} \equiv \mathcal{M}(\Lambda_j \mu_j, 0) \]  
(A.231)

\[ C_{0\mu} \equiv \mathcal{L}\left( (2 \mu_t' - 2 \theta_{\mu,t} \mu_t') \Lambda_j / T, 0 \right) \quad D_{0\mu} \equiv \mathcal{M}(\Lambda_j \mu_j, 0) \]  
(A.232)

For \( \hat{\theta}_{F,t} \), we use (A.94) to create the extended \( D \) matrix as follows:

\[ C_{0F} \equiv \mathcal{L}\left( -2 \theta_{\mu,F} V_F' / T, 0 \right) \quad D_{0F} \equiv \mathcal{M}(\Lambda_j V_F 1' \Lambda_j, 0) \]  
(A.233)

and similarly we find:

\[ C_{0FF} \equiv \mathcal{L}\left( -2 \theta_{\mu,F} V_F' \Lambda_j / T, 0 \right) \quad D_{0FF} \equiv \mathcal{M}(\Lambda_j V_F V_F', 2V_F' \Lambda_j) \]  
(A.234)

\[ C_{0\mu F} \equiv \mathcal{L}\left( (\theta_{\mu,F} \mu_t + V_F' - \theta_{\mu,F} V_F') \Lambda_j / T, 0 \right) \quad D_{0\mu F} \equiv \mathcal{M}(\Lambda_j \mu_j, \mu_j' \Lambda_j) \]  
(A.235)

Expanding and using linearity of the Theorem I matrices, we also find

\[ C_{0\mu \mu} = C_{0\mu} - \theta_{\mu,t}^2 C_{0t} + 2 \theta_{\mu,j} C_{0\mu} \quad D_{0\mu \mu} = D_{0\mu} - \theta_{\mu,t}^2 D_{0t} - 2 \theta_{\mu,t} \theta_{\mu,F} D_{0\mu} \]  
(A.236)

\[ C_{0\mu F} = C_{0\mu,F} - \theta_{\mu,t} \theta_{\mu,F} C_{0t} - \theta_{\mu,F} C_{0\mu} \quad C_{0F} = C_{0F,t} - \theta_{\mu,t} \theta_{\mu,F} C_{0\mu} \]  
(A.237)

\[ D_{0\mu \mu} = D_{0\mu} - \theta_{\mu,t}^2 D_{0t} - \theta_{\mu,F} D_{0\mu} \]
\[ C_{0F\Omega F} = C_{0FF} - \theta_{F,F}^2 C_{0F} - 20 \theta_{F,F} C_{0F} \]
\[ D_{0F\Omega F} = D_{0FF} - \theta_{F,F}^2 D_{0F} - 20 \theta_{F,F} D_{0F} \]  \hspace{1cm} (A.238)

From (A.181) for \( \gamma_1 \) along with (A.187) we may define

\[ C_{\gamma_1} = A^{-1} \sum_{t=1}^{T} Z_{t-1} C_{\theta t} \]
\[ D_{\gamma_1} = \frac{1}{T} \sum_{t=1}^{T} D_{\theta t} \]  \hspace{1cm} (A.239)

and similarly,

\[ C_{\gamma \mu} = A^{-1} \sum_{t=1}^{T} Z_{t-1} \left( \theta_{t} C_{0\mu t} + \theta_{\mu t} C_{0t} \right) \]
\[ D_{\gamma \mu} = \frac{1}{T} \sum_{t=1}^{T} \left( \theta_{t} D_{0\mu t} + \theta_{\mu t} D_{0t} \right) \]  \hspace{1cm} (A.240)

For \( \gamma_F \) and other expansions involving \( \mu_t^F \) we will need to use the Theorem I matrices \( C_t^F \) and \( D_t^F \) for \( \hat{\mu}_t^F \) as defined by (A.196)

\[ C_{\gamma F} = A^{-1} \sum_{t=1}^{T} Z_{t-1} \left( \theta_{t} C_{0F t} + \theta_{F,t} C_{0t} + \theta_{\mu t} \mu_t^F C_{0t} + \theta_{\mu t} \mu_t^F C_{0t} + \theta_{\mu t} \mu_t^F C_{0t} \right) \]
\[ D_{\gamma F} = \frac{1}{T} \sum_{t=1}^{T} \left( \theta_{t} D_{0F t} + \theta_{F,t} D_{0t} + \theta_{\mu t} \mu_t^F D_{0t} + \theta_{\mu t} \mu_t^F D_{0t} + \theta_{\mu t} \mu_t^F D_{0t} \right) \]  \hspace{1cm} (A.241)

\[ C_{\gamma \mu F} = A^{-1} \sum_{t=1}^{T} Z_{t-1} \left( C_{0\mu F t} + \mu_t^F C_{0\mu t} + \theta_{\mu \mu t} C_{0t} \right) \]
\[ D_{\gamma \mu F} = \frac{1}{T} \sum_{t=1}^{T} \left( D_{0\mu F t} + \mu_t^F D_{0\mu t} + \theta_{\mu \mu t} D_{0t} \right) \]  \hspace{1cm} (A.242)

\[ C_{\gamma FF} = A^{-1} \sum_{t=1}^{T} Z_{t-1} \left[ C_{0F\Omega F t} + 2 \mu_t^F C_{0\mu F t} + 2 \theta_{\mu \mu t} \mu_t^F C_{0t} + \left( \mu_t^F \right)^2 C_{0\mu t} + 2 \theta_{\mu \mu t} \mu_t^F C_{0t} \right] \]
\[ D_{\gamma FF} = \frac{1}{T} \sum_{t=1}^{T} \left[ D_{0F\Omega F t} + 2 \mu_t^F D_{0\mu F t} + 2 \theta_{\mu \mu t} \mu_t^F D_{0t} + \left( \mu_t^F \right)^2 D_{0\mu t} + 2 \theta_{\mu \mu t} \mu_t^F D_{0t} \right] \]  \hspace{1cm} (A.243)

We are now ready to expand \( \lambda_1, \lambda_2, \mu_p, \) and \( \sigma_p^2 \).
\[
C_{s,1} \equiv \frac{-(\mu^F - \gamma_{\mu F})C_{1} - \gamma_{1}(C^F - C_{\gamma F}) + \gamma_{F}C_{\gamma \mu} + \gamma_{\mu}C_{F}^F}{\gamma_{\mu}(\mu^F - \gamma_{\mu F}) + \gamma_{F}(\gamma_{\mu} - 1)}
\]
\[
D_{s,1} \equiv \frac{-(\mu^F - \gamma_{\mu F})D_{s,1} - \gamma_{1}(D^F - D_{\gamma F}) + \gamma_{F}D_{\gamma \mu} + \gamma_{\mu}D_{F}^F}{\gamma_{\mu}(\mu^F - \gamma_{\mu F}) + \gamma_{F}(\gamma_{\mu} - 1)}
\]
\[
C_{s,2} \equiv \frac{-\gamma_{1}C_{\gamma \mu} - (\gamma_{\mu} - 1)C_{s,1} - 2\gamma_{\mu}C_{\gamma \mu}}{\gamma_{\mu}(\mu^F - \gamma_{\mu F}) + \gamma_{F}(\gamma_{\mu} - 1)}
\]
\[
D_{s,2} \equiv \frac{-\gamma_{1}D_{s,1} - (\gamma_{\mu} - 1)D_{s,1} - 2\gamma_{\mu}D_{\gamma \mu}}{\gamma_{\mu}(\mu^F - \gamma_{\mu F}) + \gamma_{F}(\gamma_{\mu} - 1)}
\]

From (A.166) and (A.171) we find
\[
C_{\mu \mu} \equiv C_{\gamma \mu} - \gamma_{\mu}C_{s,1} - \lambda_{1}C_{\gamma \mu} - \gamma_{\mu F}C_{s,2} - \lambda_{2}C_{\gamma F}
\]
\[
D_{\mu \mu} \equiv D_{s,1} - \gamma_{\mu}D_{s,1} - \lambda_{1}D_{\gamma \mu} - \gamma_{\mu F}D_{s,2} - \lambda_{2}D_{\gamma F}
\]
\[
C_{\gamma 2} \equiv C_{s,1} + 2\gamma_{\mu \mu} \lambda_{1}C_{s,1} + \lambda_{1}^2 C_{\gamma \mu} + 2\lambda_{1} \gamma_{\mu F} C_{s,1} + 2\lambda_{2} \gamma_{\mu F} C_{s,2} + 2\lambda_{2} \gamma_{\mu F} C_{s,1} + 2\lambda_{2} \gamma_{\mu F} D_{s,1} + 2\lambda_{2} \gamma_{\mu F} D_{s,2} + 2\lambda_{2} \gamma_{\mu F} D_{s,1} + 2\lambda_{2} \gamma_{\mu F} D_{s,2} + 2\lambda_{2} \gamma_{\mu F} D_{s,1} + 2\lambda_{2} \gamma_{\mu F} D_{s,2}
\]
\[
D_{\gamma 2} \equiv D_{s,1} + 2\gamma_{\mu \mu} \lambda_{1}D_{s,1} + \lambda_{1}^2 D_{\gamma \mu} + 2\lambda_{1} \gamma_{\mu F} D_{s,1} + 2\lambda_{2} \gamma_{\mu F} D_{s,2} + 2\lambda_{2} \gamma_{\mu F} D_{s,1} + 2\lambda_{2} \gamma_{\mu F} D_{s,2} + 2\lambda_{2} \gamma_{\mu F} D_{s,1} + 2\lambda_{2} \gamma_{\mu F} D_{s,2}
\]
and, finally, we find the canonical matrices for the squared Sharpe Ratio \( S_p^2 = (\mu_p - \varphi)^2 / \sigma_p^2 \) to be
\[
C \equiv \frac{2\sigma_p^2(\mu_p - \varphi)C_{\mu p} - (\mu_p - \varphi)^2 C_{\text{op}^2}}{\sigma_p^4} \\
D \equiv \frac{2\sigma_p^2(\mu_p - \varphi)D_{\mu p} - (\mu_p - \varphi)^2 D_{\text{op}^2}}{\sigma_p^4}
\]

(A.249)

completing the proof. □

Proof of Corollary III:

The optimum zero-beta rate is \( \phi = \frac{b - b^*}{c - c^*} \). Expanding this equation we obtain:

\[
\hat{\phi} - \phi = \frac{\partial \phi}{\partial b} (\hat{b} - b) + \frac{\partial \phi}{\partial b^*} (\hat{b}^* - b^*) + \frac{\partial \phi}{\partial c} (\hat{c} - c) + \frac{\partial \phi}{\partial c^*} (\hat{c}^* - c^*) + O\left(\frac{1}{T}\right),
\]

(A.250)

where

\[
\frac{\partial \phi}{\partial b} = \frac{1}{c - c^*},
\]

\[
\frac{\partial \phi}{\partial b^*} = -\frac{1}{c - c^*},
\]

\[
\frac{\partial \phi}{\partial c} = -\frac{b - b^*}{(c - c^*)^2},
\]

\[
\frac{\partial \phi}{\partial c^*} = \frac{b - b^*}{(c - c^*)^2}.
\]

Now, \( \hat{a}, \hat{b}, \hat{c}, \hat{a}^*, \hat{b}^* \) and \( \hat{c}^* \) can be similarly expanded with canonical matrices \( C_a, D_a, C_b, D_b, C_c, D_c, C_a^*, D_a^*, C_b^*, D_b^*, C_c^*, \) and \( D_c^* \). Since \( \phi \) is a continuous differentiable function of \( \hat{a}, \hat{b}, \hat{c}, \hat{a}^*, \hat{b}^* \) and \( \hat{c}^* \), \( \hat{\phi} \) can be written in the same format. To determine the canonical matrices, based on the expansion of \( \hat{\phi} - \phi \), we obtain

\[
C_\phi = \frac{1}{c - c^*} C_b - \frac{1}{c - c^*} C_{b^*} - \frac{b - b^*}{(c - c^*)^2} C_c + \frac{b - b^*}{(c - c^*)^2} C_{c^*}.
\]

\[
D_\phi = \frac{1}{c - c^*} D_b - \frac{1}{c - c^*} D_{b^*} - \frac{b - b^*}{(c - c^*)^2} D_c + \frac{b - b^*}{(c - c^*)^2} D_{c^*}.
\]

(A.251)

Plug the optimum zero-beta rate to the difference between squared Sharpe ratios (Barillas and Shanken (2017)),

\[
BS = (a - a^*) - 2(b - b^*)\phi + (c - c^*)\phi^2.
\]

(A.252)
Expand $BS$,

\[
\hat{BS} - BS = (\hat{a} - a) - (\hat{a}^* - a^*) - 2\phi(\hat{b} - b) + 2\phi(\hat{b}^* - b^*) - 2(b - b^*)(\hat{\phi} - \phi) \\
+ \phi^2(\hat{c} - c) - \phi^2(\hat{c}^* - c^*) + 2\phi(c - c^*)(\hat{\phi} - \phi) + O(\frac{1}{T}). \tag{A.253}
\]

BS can be expanded in the same way as $\hat{\theta}$ in the Theorem I, and the canonical matrices for BS test are:

\[
C_{BS} = C_a - C_{a^*} - 2\phi(C_b - C_{b^*}) - 2(b - b^*)C\phi + \phi^2(C_c - C_{c^*}) + 2\phi(c - c^*)C\phi , \\
D_{BS} = D_a - D_{a^*} - 2\phi(D_b - D_{b^*}) - 2(b - b^*)D\phi + \phi^2(D_c - D_{c^*}) + 2\phi(c - c^*)D\phi . \tag{A.254}
\]

The Theorem I applies to this test using $C_{BS}$ and $D_{BS}$. QED.

2. Asymptotic Results with Time-varying covariances (Preliminary)

We present some analysis for how to extend the Theorem I to the case with conditional heteroscedasticity. The estimates for $C$ and $D$ for the Squared Sharpe Ratio with Constant Conditional Covariance are

\[
C'_i = \frac{2}{T} \mu'_i V^{-1} \quad \text{and} \quad D = -\frac{1}{T} \sum_{t=1}^{T} V^{-1} \mu_i \mu'_i V^{-1} = -\frac{1}{T} V^{-1} \left( \sum_{t=1}^{T} \mu_i \mu'_i \right) V^{-1}
\]

Theorem I with Time-Varying Conditional Covariance: Consider a scalar estimator of the form

\[
\hat{\theta} \equiv 0 + \sum_{t=1}^{T} C'_i (\hat{\mu}_i - \mu_i) + \sum_{t=1}^{T} tr \left[ D_i \left( \hat{V}_i - V_i \right) \right] \quad \text{where} \quad C_1, ..., C_T \quad \text{are} \quad N \times 1 \quad \text{vectors, and the} \quad D_i \quad \text{are} \quad N \times N \quad \text{matrices with} \quad D \quad \text{denoting their sum. Define the} \quad L \times N \quad \text{matrix} \quad C \equiv A^{-1} \sum_{t=1}^{T} Z^{-1}_{t-1} C'_t , \quad \text{where} \quad A \equiv \frac{1}{T} \sum_{t=1}^{T} Z^{-1}_{t-1} Z^{-1}_{t-1} , \quad \text{and let} \quad \tau_i \quad \text{denote the time-}t \quad N \times N \quad \text{matrix of residuals obtained by regressing} \quad \hat{V}_i - \hat{V} \quad \text{on} \quad Z_{t-1} \quad \text{on} \quad Z_{t-1} \quad . \quad \text{The asymptotic variance of} \quad \hat{\theta} \quad \text{may be estimated as:}
\]

\[
A\text{VAR}(\hat{\theta}) = \frac{1}{T} \sum_{t=1}^{T} \left[ \frac{1}{T} \sum_{i=1}^{T} \left( \hat{V}_{it} - \bar{V}_t \right)^2 \right] \tag{A.255}
\]
with \( v_n = Z_{i-1} \hat{C} \hat{e}_i + \hat{e}_i \hat{D} \hat{e}_i + T \text{ tr} \left( \hat{D}_i \hat{\xi}_i \right) \) and \( \bar{v}_n = \sum_{i=1}^{T} v_n / T \), where \( \hat{C} \), \( \hat{D} \), and \( \hat{\xi} \) are consistent estimates for the canonical matrices \( C \), \( D \), and \( D \) respectively.

**Proof:** We begin with an identity:

\[
\hat{\mu}' \hat{V}^{-1} \hat{\mu}_t = \left[ \mu_t + (\hat{\mu}_t - \mu_t) \right]' \left[ V + (\hat{V} - V) \right]^{-1} \left[ \mu_t + (\hat{\mu}_t - \mu_t) \right] \tag{A.256}
\]

and continue with a first-order approximation:

\[
\hat{\mu}' \hat{V}^{-1} \hat{\mu}_t \approx 2 \mu'_t V^{-1} (\hat{\mu}_t - \mu_t) + \mu'_t \left[ V + (\hat{V} - V) \right]^{-1} \mu_t \\
= 2 \mu'_t V^{-1} (\hat{\mu}_t - \mu_t) + \mu'_t \left[ I + (\hat{V} - V) V^{-1} \right]^{-1} \mu_t \\
\approx 2 \mu'_t V^{-1} (\hat{\mu}_t - \mu_t) + \mu'_t \left[ I - (\hat{V} - V) V^{-1} \right]^{-1} \mu_t \\
= 2 \mu'_t V^{-1} (\hat{\mu}_t - \mu_t) + \mu'_t \left[ V^{-1} - V^{-1} (\hat{V} - V) V^{-1} \right] \mu_t \\
= \mu_t V^{-1} \mu_t + 2 \mu'_t (\hat{\mu}_t - \mu_t) - \mu'_t (\hat{V} - V) V^{-1} \mu_t \\
= \mu_t V^{-1} \mu_t + 2 \mu'_t (\hat{\mu}_t - \mu_t) - \text{tr} \left[ V^{-1} \mu_t \mu'_t (\hat{V} - V) \right]. 
\tag{A.257}
\]

The estimated average Sharpe Ratio may therefore be written as

\[
\frac{1}{T} \sum_{t=1}^{T} \hat{\mu}' \hat{V}^{-1} \hat{\mu}_t \\
\approx \frac{1}{T} \sum_{t=1}^{T} \mu_t V^{-1} \mu_t + \frac{1}{T} \sum_{t=1}^{T} 2 \mu'_t V^{-1} (\hat{\mu}_t - \mu_t) - \frac{1}{T} \sum_{t=1}^{T} \text{tr} \left[ V^{-1} \mu_t \mu'_t (\hat{V} - V) \right] 
\tag{A.258}
\]

which is in the form of the Theorem I

\[
\hat{\theta} \approx \theta + \sum_{t=1}^{T} C_t (\hat{\mu}_t - \mu_t) + \text{tr} \left[ D (\hat{V} - V) \right] \tag{A.259}
\]

with \( C_t = \frac{2}{T} \mu'_t V^{-1} \) and \( D = -\frac{1}{T} \sum_{t=1}^{T} V^{-1} \mu_t \mu'_t V^{-1} = -\frac{1}{T} V^{-1} \left( \sum_{t=1}^{T} \mu_t \mu'_t \right) V^{-1} \).
Recall that
\[
\hat{\theta} \approx \theta + \sum_{i=1}^{T} C_i'(\hat{\mu}_i - \mu_i) + \sum_{i=1}^{T} tr\left[D_i'\left(\hat{V}_i - V_i\right)\right] \tag{A.260}
\]

We begin with the middle term on the right, and proceed as in the constant conditional variance case, to find:
\[
\sum_{i=1}^{T} C_i'(\hat{\mu}_i - \mu_i) = \sum_{i=1}^{T} C_i'\left(\frac{1}{T} \sum_{i=1}^{T} \varepsilon_i Z_{t-1}'\right) A^{-1} Z_{t-1} \tag{A.261}
\]

Using the \( L \times N \) matrix (without subscript) \( C = A^{-1} \sum_{i=1}^{T} Z_{t-1} C_i' \), and using matrix commutativity within the trace operator, this becomes
\[
\sum_{i=1}^{T} C_i'(\hat{\mu}_i - \mu_i) = tr\left[\sum_{i=1}^{T} C_i'\left(\frac{1}{T} \sum_{i=1}^{T} \varepsilon_i Z_{t-1}'\right) A^{-1} Z_{t-1} C_i'\right] \\
= tr\left[\sum_{t=1}^{T} \left(\frac{1}{T} \sum_{i=1}^{T} \varepsilon_i Z_{t-1}'\right) A^{-1} Z_{t-1} C_i'\right] \\
= tr\left[\left(\frac{1}{T} \sum_{i=1}^{T} \varepsilon_i Z_{t-1}'\right) A^{-1} \sum_{i=1}^{T} Z_{t-1} C_i'\right] = \frac{1}{T} tr\left(\sum_{i=1}^{T} \varepsilon_i Z_{t-1}' C_i\right) \\
= \frac{1}{T} \sum_{i=1}^{T} Z_{t-1}' C \varepsilon_i \tag{A.262}
\]

Because conditional constants will not contribute to the conditional variance of \( \hat{\theta} \), we may work with \( \sum_{i=1}^{T} tr\left(D_i\hat{V}_i\right) \) in place of the last term on the right for the conditional variance of \( \hat{\theta} \). We will use the following decomposition:
\[
\hat{V}_i = \hat{V} + (\hat{V}_i - \hat{V}) \equiv \frac{1}{T} \sum_{t=1}^{T} \varepsilon_i \varepsilon_i' + (\hat{V}_i - \hat{V}) \tag{A.263}
\]

We now regress \( \hat{V}_i - \hat{V} \) on \( Z \) (i.e., one bivariate regression for each matrix entry) and assume that the matrix residuals at time \( t \), denoted \( \tau_i \), are independent over time but may be correlated with the \( \varepsilon_i \). Note
that if this regression is perfect (i.e., residuals are zero) then the term $\hat{V}_t - \hat{V}$ becomes a known function of the $Z$ and does not contribute to the conditional variance; hence it is natural to concentrate on the residuals $\tau_t$ while accounting for the conditional variance contributed by $\hat{V}_t - \hat{V}$. Up to a conditional constant, for the purpose of computing the conditional variance, we may therefore represent

$$\hat{V}_t = \hat{V} + (\hat{V}_t - \hat{V})$$

using $\left(\sum_{i=1}^{T} \epsilon_i \epsilon_i'\right)/T + \tau_t$. This implies that we may, for this purpose, represent

$$\sum_{i=1}^{T} tr(D_i \hat{V}_t)$$

using

$$\sum_{i=1}^{T} tr \left[ D_i \left( \frac{1}{T} \sum_{i=1}^{T} \epsilon_i \epsilon_i' + \tau_i \right) \right] = \sum_{i=1}^{T} tr \left( D_i \frac{1}{T} \sum_{i=1}^{T} \epsilon_i \epsilon_i' \right) + \sum_{i=1}^{T} tr( D_i \tau_i )$$

where $D = \sum_{i=1}^{T} D_i$.

Putting these two terms for $\hat{\theta}$ together, for the purpose of finding its conditional variance, we may work with

$$\frac{1}{T} \sum_{j=1}^{T} Z_{-i} C \epsilon_i + \frac{1}{T} \sum_{i=1}^{T} \epsilon_i' D e_i + \sum_{i=1}^{T} tr( D_i \tau_i )$$

(A.265)

It now follows, by conditional independence of $(\epsilon_i, \tau_i)$ over time, that the conditional asymptotic variance of $\hat{\theta}$ is

$$\text{AVAR}(\hat{\theta}) = \frac{1}{T} \sum_{i=1}^{T} \text{Var} \left[ Z_{-i} C \epsilon_i + \epsilon_i' D e_i + T \ tr( D_i \tau_i ) \mid Z_0, ..., Z_{T-1} \right]$$

(A.266)

If we define $v_a = Z_{-i} \hat{C} \epsilon_i + \epsilon_i' \hat{D} e_i + T \ tr( \hat{D}_i \hat{\tau}_i )$ and $\bar{v}_a = \sum_{i=1}^{T} v_a / T$

then we may write for the unconditional variance,
\[ AVAR(\hat{\theta}) = \frac{1}{T} \sum_{t=1}^{T} \left[ \frac{1}{T} \sum_{i=1}^{T} (v_{it} - \bar{v}_{it})^2 \right] \] (A.267)

completing the proof. \[ \square \]

3. Equivalence of Two Formulations

We wish to show equivalence of two formulas for conditional portfolio weights in the presence of a time-varying conditionally riskless asset with return \( R_f = R_f(Z) \). One formula is from Theorem 3 using Equation (14) on page 977 of FS2001 for the case of no riskless asset (where we instead include the time-varying conditionally riskless asset \( R_f \) as risky asset \( n+1 \)). This formula gives the portfolio weights of the original \( n \) risky assets followed by the weight in the conditionally riskless asset as the \((n+1)\) vector:

\[
\chi' = \frac{\mathbf{1}_{n+1}' \Lambda}{\mathbf{1}_{n+1}' \mathbf{1}_{n+1}} \mu_p - \frac{\alpha_2}{\alpha_3} \mu_{n+1}' \left( \Lambda - \frac{\Lambda \mathbf{1}_{n+1} \mathbf{1}_{n+1}' \Lambda}{\mathbf{1}_{n+1}' \mathbf{1}_{n+1}} \right) \] (A.268)

where \( \mu_p \) is the target unconditional portfolio mean, \( \mu_{n+1} = \mu_{n+1}(Z) \) is the \((n+1)\) vector of conditional means whose last entry is \( R_f \), and \( V_{n+1} \) is the conditional covariance matrix of the risky assets whose final row and column are zeros,

\[
\Lambda = \Lambda(Z) = \left( \mu_{n+1} \mu_{n+1}' + V_{n+1} \right)^{-1} \quad \alpha_2 = E \left( \frac{\mathbf{1}_{n+1}' \Lambda \mu_{n+1}}{\mathbf{1}_{n+1}' \mathbf{1}_{n+1}} \right)
\]

and

\[
\alpha_3 = E \left[ \mu_{n+1}' \left( \Lambda - \frac{\Lambda \mathbf{1}_{n+1} \mathbf{1}_{n+1}' \Lambda}{\mathbf{1}_{n+1}' \mathbf{1}_{n+1}} \right) \mu_{n+1} \right] \] (A.269)

We assume that \( R_f \neq 0 \) almost surely so that the inverse \( \Lambda = \Lambda(Z) = \left( \mu_{n+1} \mu_{n+1}' + V_{n+1} \right)^{-1} \) exists.
The other formula is from Equation (35.18) of Ferson and Siegel (2015) which models the initial $n$ risky assets separately from the conditionally riskless asset $R_f$ and gives the portfolio weights of the original $n$ risky assets as the $n$ vector:

$$y' = [(c + 1)\mu_p + b - R_f'](\mu_n - R_f1_n)'Q$$  \hspace{1cm} (A.270)

where $\mu_n$ is the vector of means for the $n$ risky assets (excluding the conditionally riskless asset) and the amount invested in the conditionally riskless asset is $1 - y'1_n$ to preserve the portfolio constraint. 

$$Q = \left\{(\mu_n - R_f1_n)(\mu_n - R_f1_n)' + V_n\right\}^{-1}$$

$$b = \frac{E\left[R_f(\mu_n - R_f1_n)'Q(\mu_n - R_f1_n)\right] - E(R_f)}{E\left[(\mu_n - R_f1_n)'Q(\mu_n - R_f1_n)\right]}$$

and

$$c = \frac{1 - E\left[(\mu_n - R_f1_n)'Q(\mu_n - R_f1_n)\right]}{E\left[(\mu_n - R_f1_n)'Q(\mu_n - R_f1_n)\right]}$$  \hspace{1cm} (A.271)

Please note that, with this notation, $\mu_n$ denotes the vector of the first $n$ entries of $\mu_{n+1}$ and similarly for $V_n$ and $V_{n+1}$. That is, if we partition into the first $n$ and the last 1, we have

$$\mu_{n+1} = \begin{bmatrix} \mu_n \\ R_f \end{bmatrix}$$

$$V_{n+1} = \begin{bmatrix} V_n & 0 \\ 0 & 0 \end{bmatrix}$$  \hspace{1cm} (A.272)
Proposition: The vector of the first \( n \) entries of \( x \) is equal to \( y \). That is, \( x'_{1...n} = y' \). That is, the two methods produce the same weights.

Proof: It is sufficient to show that \( x'_{1...n} = y' \) because the final weight of \( x_{n+1} \) is the amount invested in the conditionally riskless asset and the entries of \( x_{n+1} \) sum to 1 by construction, while the portfolio defined by \( y \) is financed using this same investment in the conditionally riskless asset, thereby also achieving the portfolio constraint. We now expand \( \Lambda \) using Lemma 2 while noting that the final entry of \( \mu_{n+1} \) is \( R_f \)

\[
\Lambda = (V_{n+1} + \mu_{n+1} \mu'_{n+1})^{-1} = \left[ \begin{array}{c} V_n^{-1} & -V_n^{-1} \mu_n / R_f \\ -\mu_n' V_n^{-1} / R_f & (1 + \mu_n' V_n^{-1} \mu_n) / R_f^2 \end{array} \right]
\]

(A.273)

To find \( x' = \frac{I_{n+1}' \Lambda}{I_{n+1}' \Lambda I_{n+1}} + \mu_p - \frac{\alpha_2}{\alpha_3} \mu_{n+1}' \frac{\Lambda - \Lambda I_{n+1}' I_{n+1} \Lambda}{I_{n+1}' \Lambda I_{n+1}} \) using the expanded \( \Lambda \), we will need each of the following:

\[
I_{n+1}' \Lambda = I_{n+1}' \left[ \begin{array}{c} V_n^{-1} & -V_n^{-1} \mu_n / R_f \\ -\mu_n' V_n^{-1} / R_f & (1 + \mu_n' V_n^{-1} \mu_n) / R_f^2 \end{array} \right]
\]

\[
= I_{n+1}' V_n^{-1} - \mu_n' V_n^{-1} / R_f - I_{n+1}' V_n^{-1} \mu_n / R_f + (1 + \mu_n' V_n^{-1} \mu_n) / R_f^2 \]

\[
I_{n+1}' \Lambda I_{n+1} = \left[ I_{n+1}' V_n^{-1} - \mu_n' V_n^{-1} / R_f - I_{n+1}' V_n^{-1} \mu_n / R_f + (1 + \mu_n' V_n^{-1} \mu_n) / R_f^2 \right] I_{n+1}
\]

\[
= I_{n+1}' V_n^{-1} - \mu_n' V_n^{-1} / R_f - I_{n+1}' V_n^{-1} \mu_n / R_f + (1 + \mu_n' V_n^{-1} \mu_n) / R_f^2 \]

\[
= \frac{R_f^2 I_{n+1}' V_n^{-1} - R_f \mu_n' V_n^{-1} \mu_n / R_f + (1 + \mu_n' V_n^{-1} \mu_n)}{R_f^2}
\]

\[
= \frac{1 + (\mu_n - R_f I_n) V_n^{-1}(\mu_n - R_f I_n)}{R_f^2}
\]
\[
\begin{bmatrix}
V_n^{-1} & -V_n^{-1} \mu_n / R_f \\
-\mu_n V_n^{-1} / R_f & \left(1 + \mu_n V_n^{-1} \mu_n\right) / R_f^2
\end{bmatrix}
\]

\[
= \left[ \begin{array}{cc}
\mu_n V_n^{-1} - R_f \mu_n V_n^{-1} / R_f & -\mu_n V_n^{-1} / R_f + R_f \left(1 + \mu_n V_n^{-1} \mu_n\right) / R_f^2 \\
0 & 1 / R_f
\end{array} \right]
\]

\[
\mu_{n+1}' \Lambda = \mu_{n+1}' \left[ \begin{array}{cc}
V_n^{-1} & -V_n^{-1} \mu_n / R_f \\
-\mu_n V_n^{-1} / R_f & \left(1 + \mu_n V_n^{-1} \mu_n\right) / R_f^2
\end{array} \right]
\]

\[
= \left[ \begin{array}{cc}
\mu_n V_n^{-1} - R_f \mu_n V_n^{-1} / R_f & -\mu_n V_n^{-1} / R_f + R_f \left(1 + \mu_n V_n^{-1} \mu_n\right) / R_f^2 \\
0 & 1 / R_f
\end{array} \right] \mu_{n+1}'
\]

Using these in the formula for \( x' \) we find

\[
x' = x'_{n+1} - \frac{\mu_p - \alpha_2}{\alpha_3} \mu_{n+1}' \left( \Lambda - \frac{\Lambda \mu_{n+1}' \Lambda}{\Lambda'_{n+1} \Lambda_{n+1}} \right)
\]

\[
= \frac{\Lambda_{n+1}'}{\Lambda'_{n+1} \Lambda_{n+1}} + \frac{\mu_p - \alpha_2}{\alpha_3} \mu_{n+1}' \Lambda - \frac{\mu_p - \alpha_2}{\alpha_3} \mu_{n+1}' \Lambda_{n+1} \Lambda_{n+1}
\]

\[
= \left( \frac{1}{\alpha_3 R_f} \right) \left[ \begin{array}{cc}
1 & -\mu_n V_n^{-1} / R_f + R_f \left(1 + \mu_n V_n^{-1} \mu_n\right) / R_f^2 \\
0 & 1 / R_f
\end{array} \right] \mu_{n+1}' \Lambda
\]

\[
= \left( \frac{1}{\alpha_3 R_f} \right) \left[ \begin{array}{cc}
1 & -\mu_n V_n^{-1} / R_f + R_f \left(1 + \mu_n V_n^{-1} \mu_n\right) / R_f^2 \\
0 & 1 / R_f
\end{array} \right] \mu_{n+1}' \Lambda
\]

(A.274)
Because we know that $x'1_{n+1} = 1$ from direct calculation, we may look at the first $n$ coordinates $x_{i,n}$ of $x$, knowing that this position will be financed using a position in the conditionally risk-free asset. We find

$$x'_{i,n} = \left(1 - \frac{\mu_p - \alpha_2}{\alpha_3 R_f}\right) \frac{1^t V_n^{-1} R_f^2 - \mu'_n V_n^{-1} R_f}{1 + (\mu_n - R_f 1_n)' V_n^{-1} (\mu_n - R_f 1_n)}$$

$$= \left(\frac{\mu - \alpha_2}{\alpha_3} - R_f\right) \frac{(\mu - R_f 1_n)' V_n^{-1}}{1 + (\mu_n - R_f 1_n)' V_n^{-1} (\mu_n - R_f 1_n)}$$

(A.267)

Now use Lemma 4 to expand

$$Q = \left\{V_n + (\mu_n - R_f 1_n)(\mu_n - R_f 1_n)'\right\}^{-1}$$

$$= V_n^{-1} - V_n^{-1} (\mu_n - R_f 1_n)(\mu_n - R_f 1_n)' V_n^{-1} \left(\mu_n - R_f 1_n\right)$$

(A.277)

from which we find

$$y' = \left[(c + 1)\mu_p + b - R_f\right] (\mu_n - R_f 1_n)' Q$$

$$= \left[(c + 1)\mu_p + b - R_f\right] (\mu_n - R_f 1_n)' \left(V_n^{-1} - V_n^{-1} (\mu_n - R_f 1_n)(\mu_n - R_f 1_n)' V_n^{-1} \left(\mu_n - R_f 1_n\right)\right)$$

$$= \left[(c + 1)\mu_p + b - R_f\right] \left(V_n^{-1} - \frac{(\mu_n - R_f 1_n)' V_n^{-1} (\mu_n - R_f 1_n)(\mu_n - R_f 1_n)' V_n^{-1} \left(\mu_n - R_f 1_n\right)}{1 + (\mu_n - R_f 1_n)' V_n^{-1} (\mu_n - R_f 1_n)}\right)$$

$$= \left[(c + 1)\mu_p + b - R_f\right] \left(1 - \frac{(\mu_n - R_f 1_n)' V_n^{-1} (\mu_n - R_f 1_n)}{1 + (\mu_n - R_f 1_n)' V_n^{-1} (\mu_n - R_f 1_n)}\right)$$

$$= \left[(c + 1)\mu_p + b - R_f\right] \frac{(\mu_n - R_f 1_n)' V_n^{-1}}{1 + (\mu_n - R_f 1_n)' V_n^{-1} (\mu_n - R_f 1_n)}$$
Comparing $x'_1 \ldots x'_n$ to $y'$, we see that they will be equal provided we show that

$$\frac{\mu_p - \alpha_2}{\alpha_3} - R_j = (c + 1)\mu_p + b - R_j$$  \hspace{1cm} (A.279)

or, more simply, that

$$\frac{\mu_p - \alpha_2}{\alpha_3} = (c + 1)\mu_p + b$$  \hspace{1cm} (A.280)

We next show that the multiples of $\mu_p$ are identical, that is: $\alpha_3 = 1 / (1 + c)$. We find

$$\alpha_3 = E \left[ \mu'_{n+1} \left( \Lambda - \frac{\Lambda \mathbf{1}_{n+1} \mathbf{1}'_{n+1} \Lambda}{\mathbf{1}'_{n+1} \Lambda \mathbf{1}_{n+1}} \right) \mu_{n+1} \right] = E \left( \mu'_{n+1} \Lambda \mu_{n+1} - \frac{\mu'_{n+1} \Lambda \mathbf{1}_{n+1} \mathbf{1}'_{n+1} \Lambda \mu_{n+1}}{\mathbf{1}'_{n+1} \Lambda \mathbf{1}_{n+1}} \right)$$

$$= E \left( 1 - \frac{1}{R_j^2 \mathbf{1}'_{n+1} \Lambda \mathbf{1}_{n+1}} \right) = E \left( 1 - \frac{R_j^2}{R_j^2 \left[ 1 + \left( \mu_n - R_j \mathbf{1}_n \right)' V_n^{-1} \left( \mu_n - R_j \mathbf{1}_n \right) \right] \right)$$  \hspace{1cm} (A.281)

and

$$\frac{1}{1 + c} = 1 - E \left[ \left( \mu_n - R_j \mathbf{1}_n \right)' Q \left( \mu_n - R_j \mathbf{1}_n \right) \right]$$

$$= E \left[ \left( \mu_n - R_j \mathbf{1}_n \right)' Q \left( \mu_n - R_j \mathbf{1}_n \right) \right]$$  \hspace{1cm} (A.282)

From the derivation of $y'$, we know that
\[
\left( \mu_n - R_f \mathbf{1}_n \right) Q = \frac{\left( \mu_n - R_f \mathbf{1}_n \right)' V_n^{-1}}{1 + \left( \mu_n - R_f \mathbf{1}_n \right)' V_n^{-1} \left( \mu_n - R_f \mathbf{1}_n \right)}
\]

(A.283)

which implies that

\[
\frac{1}{1 + c} = E \left[ \left( \mu_n - R_f \mathbf{1}_n \right)' Q \left( \mu_n - R_f \mathbf{1}_n \right) \right] = E \left[ \frac{\left( \mu_n - R_f \mathbf{1}_n \right)' V_n^{-1}}{1 + \left( \mu_n - R_f \mathbf{1}_n \right)' V_n^{-1} \left( \mu_n - R_f \mathbf{1}_n \right)} \left( \mu_n - R_f \mathbf{1}_n \right) \right]
\]

\[
= E \left[ \frac{1 + \left( \mu_n - R_f \mathbf{1}_n \right)' V_n^{-1} \left( \mu_n - R_f \mathbf{1}_n \right) - 1}{1 + \left( \mu_n - R_f \mathbf{1}_n \right)' V_n^{-1} \left( \mu_n - R_f \mathbf{1}_n \right)} \right] = E \left( 1 - \frac{1}{1 + \left( \mu_n - R_f \mathbf{1}_n \right)' V_n^{-1} \left( \mu_n - R_f \mathbf{1}_n \right)} \right) = \alpha_3
\]

(A.284)

showing that the coefficients multiplying \( \mu_p \) are identical. It remains only to show that the remaining terms are equal, that is, that \( b = -\alpha_2 / \alpha_3 \). For this we need
\[ b = \frac{E \left[ R_f \left( \mu_n - R_f 1_n \right)'Q(\mu_n - R_f 1_n) \right]}{E \left[ (\mu_n - R_f 1_n)'Q(\mu_n - R_f 1_n) \right]} - E(R_f) \]

\[ = \frac{E \left[ R_f \frac{(\mu_n - R_f 1_n)'V_n^{-1}(\mu_n - R_f 1_n)}{1 + (\mu_n - R_f 1_n)'V_n^{-1}(\mu_n - R_f 1_n)} \right]}{E \left[ \frac{(\mu_n - R_f 1_n)'V_n^{-1}(\mu_n - R_f 1_n)}{1 + (\mu_n - R_f 1_n)'V_n^{-1}(\mu_n - R_f 1_n)} \right]} - E(R_f) \]

\[ = \frac{E \left[ \frac{R_f}{1 + (\mu_n - R_f 1_n)'V_n^{-1}(\mu_n - R_f 1_n)} \right]}{E \left[ \frac{R_f}{1 + (\mu_n - R_f 1_n)'V_n^{-1}(\mu_n - R_f 1_n)} \right]} - E(R_f) \]

\[ = \frac{E \left[ \frac{1}{1 + (\mu_n - R_f 1_n)'V_n^{-1}(\mu_n - R_f 1_n)} \right]}{E \left[ \frac{1}{1 + (\mu_n - R_f 1_n)'V_n^{-1}(\mu_n - R_f 1_n)} \right]} - E(R_f) \]
\[ \alpha_2 = E \left( \frac{1'}{n+1} \Lambda \mu_{n+1} \right) = E \left( \frac{1/R_f}{1 + \left( \mu_n - R_f 1_n \right)' V_n^{-1} \left( \mu_n - R_f 1_n \right)} \right) \]

\[ = E \left( \frac{R_f}{1 + \left( \mu_n - R_f 1_n \right)' V_n^{-1} \left( \mu_n - R_f 1_n \right)} \right) \]  \hspace{1cm} (A.285)

Putting these together with the \( \alpha_3 \) derived earlier, we find that

\[ \frac{-\alpha_2}{\alpha_3} = -\frac{E \left( \frac{R_f}{1 + \left( \mu_n - R_f 1_n \right)' V_n^{-1} \left( \mu_n - R_f 1_n \right)} \right)}{E \left( \frac{1}{1 + \left( \mu_n - R_f 1_n \right)' V_n^{-1} \left( \mu_n - R_f 1_n \right)} \right)} = b \]  \hspace{1cm} (A.286)

completing the proof. \( \Box \)

Lemma 1: Let \( u \) be an \( n + 1 \) vector with its last entry \( u_{n+1} \neq 0 \), and let \( u_{1...n} = (u_1,...,u_n)' \). Then

\[ \left( \begin{bmatrix} I_n & 0 \\ 0 & 0 \end{bmatrix} + uu' \right)^{-1} = \begin{bmatrix} I_n & -u_{1...n} / u_{n+1} \\ -u_{1...n} / u_{n+1} & \left( 1 + \left\| u_{1...n} \right\|^2 \right) / u_{n+1}^2 \end{bmatrix} \]  \hspace{1cm} (A.287)

Proof:
\[
\begin{bmatrix}
I_n & 0 \\
0 & 0
\end{bmatrix} + uu' = 
\begin{bmatrix}
I_n & -u_{1...n} / u_{n+1} \\
-u_{1...n} / u_{n+1} & (1 + \|u_{1...n}\|^2) / u_{n+1}^2
\end{bmatrix}
\]

\[
= \begin{bmatrix}
I_n & 0 \\
0 & 0
\end{bmatrix} + \begin{bmatrix}
u_{1...n}u_{1...n} & u_{1...n}u_{n+1} \\
u_{1...n}u_{n+1} & u_{n+1}^2
\end{bmatrix} = 
\begin{bmatrix}
I_n & -u_{1...n} / u_{n+1} \\
-u_{1...n} / u_{n+1} & (1 + \|u_{1...n}\|^2) / u_{n+1}^2
\end{bmatrix}
\]  

(A.288)

The top left entry of this matrix product is

\[
(I_n + u_{1...n}u'_{1...n})I_n + u_{1...n}u_{n+1}(-u'_{1...n} / u_{n+1}) = I_n
\]

(A.289)

The top right entry is

\[
(I_n + u_{1...n}u'_{1...n})(-u_{1...n} / u_{n+1}) + u_{1...n}u_{n+1}(1 + \|u_{1...n}\|^2) / u_{n+1}^2
\]

\[
= I_n(-u_{1...n} / u_{n+1}) + u_{1...n}u_{1...n}(-u_{1...n} / u_{n+1}) + u_{1...n}u_{n+1} / u_{n+1}^2 + u_{1...n}u_{n+1} \|u_{1...n}\|^2 / u_{n+1}^2
\]

\[
= -u_{1...n} / u_{n+1} - u_{1...n} \|u_{1...n}\|^2 / u_{n+1} + u_{1...n} / u_{n+1} + u_{1...n} \|u_{1...n}\|^2 / u_{n+1} = 0_n
\]

(A.290)

The bottom left entry is

\[
u'_{1...n}u_{n+1}I_n + u_{n+1}^2(-u'_{1...n} / u_{n+1}) = u'_{1...n}u_{n+1} - u_{n+1}u'_{1...n} = 0'
\]

(A.291)

Finally, the bottom right entry is

\[
u'_{1...n}u_{n+1}(-u_{1...n} / u_{n+1}) + u_{n+1}^2(1 + \|u_{1...n}\|^2) / u_{n+1}^2
\]

\[
= -u_{1...n}u_{1...n} + u_{n+1}^2 / u_{n+1} + u_{n+1} \|u_{1...n}\|^2 / u_{n+1}^2
\]

\[
= -\|u_{1...n}\|^2 + 1 + \|u_{1...n}\|^2 = 1
\]

(A.292)

Therefore

\[
\begin{bmatrix}
I_n & 0 \\
0 & 0
\end{bmatrix} + uu' = 
\begin{bmatrix}
I_n & -u_{1...n} / u_{n+1} \\
-u_{1...n} / u_{n+1} & (1 + \|u_{1...n}\|^2) / u_{n+1}^2
\end{bmatrix} = 
\begin{bmatrix}
I_n & 0' \\
0' & 1
\end{bmatrix} = I_{n+1}
\]

(A.293)
completing the proof. □

Lemma 2: Let \( u \) be an \( n+1 \) vector with its last entry \( u_{n+1} \neq 0 \), and let \( u_{1\ldots n} = (u_1, \ldots, u_n)' \). Let \( V_n \) be an \( n \times n \) symmetric positive definite matrix. Then

\[
\left( \begin{bmatrix} V_n & 0 \\ 0 & 0 \end{bmatrix} + uu' \right)^{-1} = \begin{bmatrix} V_n^{-1} \\ -V_n^{-1}u_{1\ldots n}/u_{n+1} \end{bmatrix} + \begin{bmatrix} -u_{1\ldots n}V_n^{-1}/u_{n+1} \\ 1 + u_{1\ldots n}'V_n^{-1}u_{1\ldots n}/u_{n+1}^2 \end{bmatrix}/u_{n+1}^2 \] (A.294)

Proof: Let \( V_n = AA' \) be the Cholesky decomposition. Lemma 1 implies that

\[
\left( \begin{bmatrix} I_n & 0 \\ 0 & 0 \end{bmatrix} + A^{-1} 0 \\ 0 & 0 \end{bmatrix} u \begin{bmatrix} A^{-1} & 0 \\ 0 & 1 \end{bmatrix} \right)^{-1} = \begin{bmatrix} I_n \\ -A^{-1}u_{1\ldots n}/u_{n+1} \end{bmatrix} + \begin{bmatrix} -A^{-1}u_{1\ldots n}/u_{n+1} \\ 1 + \|A^{-1}u_{1\ldots n}\|^2/u_{n+1}^2 \end{bmatrix}/u_{n+1}^2 \] (A.295)

Pre- and post-multiplying the left-hand side, we find

\[
\left( \begin{bmatrix} (A^{-1})' & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} I_n & 0 \\ 0 & 0 \end{bmatrix} + A^{-1} 0 \\ 0 & 1 \end{bmatrix} u \begin{bmatrix} A^{-1} & 0 \\ 0 & 1 \end{bmatrix} \right)^{-1} = \begin{bmatrix} (A^{-1})' & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} A^{-1} & 0 \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} AA' & 0 \\ 0 & 0 \end{bmatrix} + uu' = \left( \begin{bmatrix} V_n & 0 \\ 0 & 0 \end{bmatrix} + uu' \right)^{-1} \] (A.296)

Multiplying similarly for the right-hand side, we find
Equating these expressions completes the proof. □

Lemma 3: Let \( u \) be a vector. Then

\[
(I + uu')^{-1} = I - \frac{uu'}{1 + u'u} \quad (A.298)
\]

Proof: We multiply:

\[
(I + uu') \left( I - \frac{uu'}{1 + u'u} \right) = \left( I - \frac{uu'}{1 + u'u} \right) + uu' \left( I - \frac{uu'}{1 + u'u} \right) = I \quad (A.299)
\]

completing the proof. □

Lemma 4: Let \( u \) be a vector and let \( V \) be a symmetric positive-definite matrix. Then

\[
(V + uu')^{-1} = V^{-1} - \frac{V^{-1}uu'V^{-1}}{1 + u'V^{-1}u} \quad (A.300)
\]

Proof: Let \( V = AA' \) be the Cholesky decomposition. Then we have:
\[ (V + uu')^{-1} = (AA' + uu')^{-1} = (A')^{-1} \left[ I + \left( A^{-1} u \right) \left( A^{-1} u \right) \right]^{-1} A^{-1} \]  (A.301)

Applying Lemma 3, we find

\[
(V + uu')^{-1} = (A')^{-1} \left( I - \frac{\left( A^{-1} u \right) \left( A^{-1} u \right)'}{1 + \left( A^{-1} u \right)' \left( A^{-1} u \right)} \right) A^{-1}
\]

\[
= (A')^{-1} A^{-1} - \frac{(A')^{-1} \left( A^{-1} u \right) \left( A^{-1} u \right)'}{1 + \left( A^{-1} u \right)' \left( A^{-1} u \right)} = \left( AA' \right)^{-1} - \frac{(A')^{-1} A^{-1} uu' \left( A^{-1} \right)'}{1 + u' \left( A^{-1} \right)' \left( A^{-1} \right) u}
\]

\[
= \left( AA' \right)^{-1} - \frac{uu' \left( AA' \right)^{-1}}{1 + u' \left( AA' \right)^{-1} u} = V^{-1} - \frac{V^{-1} uu' V^{-1}}{1 + u' V^{-1} u}
\]

completing the proof. \( \square \)

4 The Impacts of Dynamic Trading

We dig into the sources of the Sharpe ratio improvements from dynamic trading of the factors in the FF3, FF5 and Q4 models. The largest squared Sharpe ratio available in each design is \( S_{ue(r,f)}^2 \), the squared Sharpe ratio attainable by dynamically trading the test assets and a model’s factors. The maximum difference between any two squared Sharpe ratios in each design is \( [S_{ue(r,f)}^2 - S_{fix(f)}^2] \). Two alternative decompositions of this maximum Sharpe ratio difference are shown:

Max \( S^2 \) Difference = \( [S_{ue(r,f)}^2 - S_{fix(f)}^2] \)

\[
= [S_{fix(f)}^2 - S_{fix(f)}^2] + [S_{ue(r,f)}^2 - S_{fix(r,f)}^2]  \]  (A.303)

\[
= [S_{ue(r,f)}^2 - S_{ue(f)}^2] + [S_{ue(f)}^2 - S_{fix(f)}^2].
\]
The first decomposition shows the sum of the squared Sharpe ratio differences associated with a classical fixed-weight factor model test, plus a measure of how trading dynamically expands the mean variance boundary of the factors and the test assets combined. The second decomposition shows the squared Sharpe ratio differences associated with a dynamic model test, plus a measure of how trading dynamically increases the squared Sharpe ratio available from the factors alone.

Table A.1 shows that all of the pieces of the decompositions are statistically significant, except in the Q4 model, where the maximum squared Sharpe ratio of the models factors is not significantly improved by dynamic trading, and in the FF5 model with the 25 Investment x productivity portfolios. The fixed weight tests do reject the FF3, FF5 and Q4 models, except in the 25 investment-productivity portfolios, where neither the FF5 nor the Q4 models is rejected.

The point estimates of the first decomposition indicate that the ability of dynamic trading to improve the squared Sharpe ratio of the combined model factors and test assets is the largest component, accounting for 40-55% of the total maximum squared Sharpe ratio. The additional squared Sharpe ratio attributed to dynamic trading is economically significant. Previous studies find economically large benefits to optimally trading with lagged instruments in out of sample analyses (see Abhyankar, Basu and Stremme (2005) and Chiang (2016) for example). This is different from simple predictive regressions, where the variables enter linearly and the out-of sample performance is poor. Here the variables enter nonlinearly through the optimal dynamic strategy. In the first decomposition, a fixed weight portfolio of the factors accounts for 12-37% of the total squared Sharpe ratio, with the largest value holding in the Q4 model, while the contribution associated with the classical fixed factor model test is 23-34% of the total. Thus, tests of the factor models that do not allow for dynamic trading miss a large fraction of the story.

The point estimates of the second decomposition indicate that the dynamic model test accounts for 60-80% of the total squared Sharpe ratio, and thus the quadratic utility attained by a UE portfolio of the factors and test assets combined. None of the models does a good job of capturing the maximum quadratic utility implied by a dynamic strategy. A fixed weight portfolio of the factors accounts for 12-37% of the total squared
Sharpe ratio. The contribution made by expanding the mean variance boundary of the factors with dynamic trading is smaller and varies across the models. This contributes 16% of the squared Sharpe ratio in the FF5 model, 9% in the FF3, but only 2% in the Q4 model.
Table A.1: Decomposing Sharpe Ratio Improvements

**Panel A: 25 size-value portfolios**

<table>
<thead>
<tr>
<th></th>
<th>[Ue(r,f) - fix(r,f)]</th>
<th>+</th>
<th>[fix(r,f)-fix(f)]</th>
<th>=</th>
<th>[Ue(r,f)-fix(f)]</th>
<th>=</th>
<th>[Ue(r,f) - ue(f)]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FF3</strong></td>
<td>BS test value</td>
<td>0.18</td>
<td>0.11</td>
<td>0.29</td>
<td>0.26</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>t-statistics</td>
<td>7.02</td>
<td>3.38</td>
<td>6.94</td>
<td>6.35</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>FF5</strong></td>
<td>BS test value</td>
<td>0.20</td>
<td>0.09</td>
<td>0.29</td>
<td>0.23</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>t-statistics</td>
<td>7.41</td>
<td>2.82</td>
<td>6.87</td>
<td>5.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Q4</strong></td>
<td>BS test value</td>
<td>0.19</td>
<td>0.11</td>
<td>0.30</td>
<td>0.28</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>t-statistics</td>
<td>6.74</td>
<td>3.07</td>
<td>6.56</td>
<td>6.10</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Panel B: 25 investment-productivity portfolios**

<table>
<thead>
<tr>
<th></th>
<th>BS test value</th>
<th>0.15</th>
<th>0.06</th>
<th>0.20</th>
<th>0.17</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>t-statistics</td>
<td>6.11</td>
<td>2.28</td>
<td>5.88</td>
<td>5.17</td>
</tr>
<tr>
<td><strong>FF5</strong></td>
<td>BS test value</td>
<td>0.16</td>
<td>0.01</td>
<td>0.17</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>t-statistics</td>
<td>6.37</td>
<td>0.56</td>
<td>5.35</td>
<td>3.78</td>
</tr>
<tr>
<td><strong>Q4</strong></td>
<td>BS test value</td>
<td>0.14</td>
<td>0.02</td>
<td>0.16</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>t-statistics</td>
<td>5.41</td>
<td>0.82</td>
<td>4.58</td>
<td>4.15</td>
</tr>
</tbody>
</table>
Panel C: 32 size-investment-productivity portfolios

<table>
<thead>
<tr>
<th></th>
<th>[Ue(r,f) - fix(r,f)]</th>
<th>+</th>
<th>[fix(r,f)-fix(f)]</th>
<th>=</th>
<th>[Ue(r,f)-fix(f)]</th>
<th>=</th>
<th>[Ue(r,f) - ue(f)]</th>
</tr>
</thead>
<tbody>
<tr>
<td>FF3</td>
<td>BS test value</td>
<td>0.18</td>
<td>0.17</td>
<td>0.35</td>
<td>0.32</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>t-statistics</td>
<td>6.90</td>
<td>4.04</td>
<td>7.06</td>
<td>6.55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FF5</td>
<td>BS test value</td>
<td>0.19</td>
<td>0.12</td>
<td>0.31</td>
<td>0.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>t-statistics</td>
<td>7.16</td>
<td>3.17</td>
<td>6.66</td>
<td>5.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q4</td>
<td>BS test value</td>
<td>0.19</td>
<td>0.09</td>
<td>0.28</td>
<td>0.27</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>t-statistics</td>
<td>6.73</td>
<td>2.38</td>
<td>5.98</td>
<td>5.63</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel D: 49 industry portfolios

<table>
<thead>
<tr>
<th></th>
<th>FF3</th>
<th>BS test value</th>
<th>0.27</th>
<th>0.20</th>
<th>0.47</th>
<th>0.44</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>t-statistics</td>
<td>8.59</td>
<td>4.77</td>
<td>8.78</td>
<td>8.28</td>
<td></td>
</tr>
<tr>
<td></td>
<td>FF5</td>
<td>BS test value</td>
<td>0.29</td>
<td>0.25</td>
<td>0.55</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td>t-statistics</td>
<td>8.72</td>
<td>5.28</td>
<td>8.77</td>
<td>8.04</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q4</td>
<td>BS test value</td>
<td>0.30</td>
<td>0.15</td>
<td>0.44</td>
<td>0.43</td>
</tr>
<tr>
<td></td>
<td>t-statistics</td>
<td>8.01</td>
<td>3.23</td>
<td>7.13</td>
<td>7.13</td>
<td></td>
</tr>
</tbody>
</table>

Squared Sharpe ratio differences and their t-ratios are shown. Ue refers to a portfolio of the indicated assets that trades optimally with the conditioning information. The test asset portfolios are denoted by r, and the factors by f.
5. Results for Alternative Lagged Instruments

We show results here using a set of “classical” lagged instruments Z including: (i) the lagged value of a one-month Treasury bill yield, (ii) the dividend yield of the market index; (iii) the spread between Moody's Baa and Aaa corporate bond yields; (iv) the spread between ten-year and one-year constant maturity Treasury bond yields. We choose these variables because they have been among the mostly commonly used in the asset pricing literature.

We also examine a set of more “modern” lagged instruments. Goyal Welch and Zafirov (GWZ, 2022) dismiss most of the 46 predictors they examine because their predictive ability either becomes insignificant in data extended to 2020, has different signs in subsamples or has poor step ahead (OOS) predictive ability. GWZ find that among our “classical” predictors only the short term tbill rate survives all of their criteria for a predictor. We pick a set of monthly modern predictors to serve as a robustness check on our empirical findings. We pick monthly variables which have data available for our sample and, even if they fail on some of the three GWZ criteria, manage to satisfy two. GWZ must have been able to roughly replicate the in-sample performance found in the original study, the OOS R-squared cannot be negative, and the extended sample t-ratio must be of the same sign as in the original study, but not necessarily statistically significant. We use the following as our modern predictors:

1. Illiq is from Chen, Eaton, and Paye (2018). This is the log of the number of zero returns, measuring stock market illiquidity. The series has structural break adjustments for tick-size reductions in 1997 and 2001, found by regressing the series on dummy variables equal to 1 after the tick-size reductions, and 0 otherwise, then taking the residuals.

2. New durables orders from Jones and Tuzel (2013) is the ratio of new orders to shipments of durable goods, obtained from the Census Bureau.

3. Technical indicators is from Neely, Rapach, Tu, and Zhou (2014). This is the first principal component of 14 technical indicators, mainly versions of moving price averages, momentum, and dollar trading volume.

4. Average Correlation is from Pollet and Wilson (2010). This is the average correlation between daily stock returns among the 500 largest stocks (by capitalization). The daily pairwise correlations of stock returns are multiplied by the product of both stock’s weights relative to total sample market capitalization, then summed.

Tables of results with these alternative instruments follow.

6. Non-traded Factors

When the factors in a model are not traded assets, or the model is fully conditional, mimicking portfolios must be found. With dynamic trading Ferson and Siegel (2009) show that given m satisfies the
pricing Equation (3), then a portfolio that is maximum correlation to \( m \) with respect to lagged information \( Z \), must be UE with respect to \( Z \). Tests compare the squared Sharpe ratio of the maximum correlation portfolio with \( S^2_{ue}(r) \). Supressing the time subscripts, we state a definition.

**Definition.** A portfolio \( R_p \) is maximum correlation for a random variable, \( m \), with respect to lagged conditioning information \( Z \), iff:

\[
\rho^2 (R_p, m) \geq \rho^2 [w'(Z)R, m] \quad \forall w(Z) : w'(Z)1 = 1,
\]

where \( \rho^2(\cdot, \cdot) \) is the squared unconditional correlation coefficient and we restrict to functions \( w(.) \) for which the correlation exists.

Ferson, Siegel and Xu (2006) present solutions for maximum correlation portfolio weights with respect to conditioning information. Tests with nontraded factors and fully conditional models depend on the choice of test assets from which the mimicking portfolios are estimated. Tests with dynamic trading depend on the choice of the lagged instruments, and we compare several specifications. In models with non-traded factors and conditioning information, a portfolio with maximum correlation (with respect to \( Z \)) to the SDF is hypothesized to be UE. Tests compare the squared Sharpe ratio of the mimicking portfolio to the maximum squared Sharpe ratio of the test assets.

Let \( R^*_i \equiv (R'_i, F_i)' = (R'^i, ..., R'^n, F_i)' \), \( \mu^*_i = (\mu'^i, \mu'^F) = (\mu'^i, ..., \mu'^n, \mu'^F) \)' , and \( \varepsilon_i^* = (\varepsilon_i', \varepsilon_i^F)' = (\varepsilon'^i, ..., \varepsilon'^n, \varepsilon'^F)' \). Define the \( k \times (N+1) \) matrix \( \delta^* = [\delta, \delta^F] \) where \( \delta^F \) contains the \( L \) regression coefficients for \( F \) given \( Z \). Define the \((N + 1) \times (N + 1)\) conditional covariance matrix

\[
V^* \equiv \begin{bmatrix}
V & V_F \\
V'_F & \sigma^2_{F|Z}
\end{bmatrix}
\]
where $V$ is the same $N \times N$ conditional covariance matrix as before, $V_F \equiv \text{Cov}(R_t F_t | Z_{t-1})$ is $N \times 1$, and
\[
\sigma_{F|Z}^2 \equiv \text{Var}(F_t | Z_{t-1})
\]
is a scalar. We assume that the unobserved $e_t^*$ are independent and identically distributed, with mean zero and covariance matrix $V^*$. We note that
\[
E(R_t F_t | Z_{t-1}) = \text{Cov}(R_t F_t | Z_{t-1}) + \mu^*_F = V_F + \mu^*_F
\]

Using Ferson, Siegel, and Xu (2006) Equation (6) the maximal correlation portfolio weight is
\[
w_t = \frac{1' \Lambda_t}{1' \Lambda_t 1} - \left[ \hat{\lambda}_1^* \mu_t^* + \hat{\lambda}_2^* E(F_t | Z_{t-1}) \right] \left( \Lambda_t - \frac{1' \Lambda_t 1}{1' \Lambda_t 1} \right)
\]

With $\Lambda_t \equiv (\mu_t^* + V)^{-1}$, $\Omega_t \equiv \Lambda_t - \frac{\Lambda_t 11' \Lambda_t}{1' \Lambda_t 1}$, $\lambda_1 \equiv -\frac{\gamma_1 (\mu^F - \gamma_{\mu F}) + \gamma_{\mu} F_F}{\gamma_{\mu} (\mu^F - \gamma_{\mu F}) + F_F (\gamma_{\mu F} - 1)}$ and $\lambda_2 \equiv \frac{-\gamma_1 (\gamma_{\mu F} - 1) - \gamma_{\mu}^2}{\gamma_{\mu} (\mu^F - \gamma_{\mu F}) + F_F (\gamma_{\mu F} - 1)}$. The parameters $\mu_F$, $\gamma_1$, $\gamma_{\mu}$, $F_F$, $\gamma_{\mu F}$ and $\gamma_{\mu \mu}$ and their estimates are presented with the proof.

**Corollary**: The asymptotic variance of the estimated squared Sharpe Ratio of the portfolio $R_p$ having maximal correlation with a given scalar factor $F$ with respect to lagged information $Z$ may be obtained using the Theorem I together with canonical matrices

\[
C \equiv \frac{2 \sigma_p^2 (\mu_p - \varphi) C_{\mu p} - (\mu_p - \varphi)^2 C_{\varphi p}}{\sigma_p^4}
\]
\[
D \equiv \frac{2 \sigma_p^2 (\mu_p - \varphi) D_{\mu p} - (\mu_p - \varphi)^2 D_{\varphi p}}{\sigma_p^4}
\]

where $\varphi$ is the given zero-beta rate, $\mu_p$ and $\sigma_p^2$ are the mean and variance of the maximal correlation portfolio, and expressions for the matrices $C_{\mu p}$, $C_{\varphi p}$, $D_{\mu p}$, and $D_{\varphi p}$ are provided with the proof. The Corollary applies to any general factor, $F$. It does not impose the assumption that the factor is the SDF.
Our data for nontraded factors follow Chen, Roll and Ross (1986). Their factors include a monthly growth in industrial production with a one-month lead, a change in expected inflation following Fama and Gibbons (1984), a corresponding unexpected inflation based on the US CPI. They also examine two traded factors: the return difference between Baa corporate bond index and the long term government bond return and the difference between the long term bond and the short term bill return. We compute a real consumption growth using the personal consumption expenditure (PCE in Table 2.3.5U) and price index data (Table 2.3.4U) from Bureau of Economic Analysis. These monthly data start from January of 1959 until December 2020. We also use a non-traded broker-dealer leverage factor, which Adrian, Etula and Muir (2014) propose as a single-factor model, the data graciously provided on Tyler Muir’s web site.

In Table 8 we estimate mimicking portfolio weights in each simulation trial to capture the effects of estimation error. The “true” values of the mimicking portfolio squared Sharpe ratios are computed using simulations with 1,000 times as many time series as in the original data. Bias adjustment uses the JK (1980) adjustment. The largest true squared Sharpe ratios for the individual mimicking portfolios are for the leverage and industrial production factors, followed by consumption growth. These range from 1.5 to 2.5% per month. The goal of the dynamic trading is to increase the squared correlation with the mimicked factor. The absolute correlation with the consumption factor increases to 0.28, versus 0.26 with fixed weights. For industrial production it rises from 0.24 to 0.34. For unexpected inflation it rises from 0.29 to 0.36.

Because mimicking portfolios are not formed to maximize their sample Sharpe ratios, they should not be as biased as maximum Sharpe ratio portfolios. The Sharpe ratios of the fixed-weight consumption and industrial production portfolios actually have a small downward bias. Other mimicking portfolios show an upward bias, especially when the true squared Sharpe ratio is close to zero (unexpected inflation).
The JK (1980) adjustment works in the right direction in these cases, because the adjustment shrinks the estimated Sharpe ratio towards zero. No bias adjustment is available for the dynamically trading mimicking portfolios.

Near the bottom of the table we combine the three nontraded Chen, Roll and Ross (1986) mimicking portfolios (CRR3) to find their maximum squared Sharpe ratio, which is just over 3.0% with fixed weights and 4.6% with time-varying weights. The mimicking portfolios are formed with fixed weights and combined into portfolios with time-varying weights.

The right hand columns of Table 8 report the empirical standard deviations of the mimicking portfolio squared Sharpe ratios, taken across the 1,000 simulation trials, and the average result from the Theorem I, Corollary III. The asymptotics do a pretty good job of predicting the simulated standard errors for most of the mimicking portfolios, getting within 10% in 40% of the cases, but they greatly overstate the sampling variability when the true Sharpe ratios are close to zero (unexpected inflation).
Table 1. Summary Statistics of factors and lagged conditioning information

This table contains summary statistics for our sample of factors from the French data library (February 1959 to December 2020), the Q factors from Hou, Xue and Zhang (2015), monthly consumption growth data (from January 1959 to Dec 2020), and the Broker-dealer leverage factor (from Adrian, Etula and Muir 2014) (from December 1970 to November 2018). The coefficient, $\rho_1$, is the first order autocorrelation (after stochastic detrending in the case of the lagged instruments). Squared SR is the squared Sharpe ratio, where the zero-beta rate is the average Treasury bill rate, equal to 0.39 percent per month. The R-square is obtained by regressing market excess return on conditioning information. Returns, yields and yield spreads are measured as percent per month.

<table>
<thead>
<tr>
<th>Model Factors</th>
<th>Mean</th>
<th>Std</th>
<th>AR(1)</th>
<th>Squared SR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market-risk free</td>
<td>0.57</td>
<td>4.43</td>
<td>0.066</td>
<td>0.016</td>
</tr>
<tr>
<td>SMB</td>
<td>0.2</td>
<td>2.96</td>
<td>0.064</td>
<td>0.005</td>
</tr>
<tr>
<td>HML</td>
<td>0.26</td>
<td>2.81</td>
<td>0.179</td>
<td>0.008</td>
</tr>
<tr>
<td>RMW</td>
<td>0.23</td>
<td>2.09</td>
<td>0.149</td>
<td>0.012</td>
</tr>
<tr>
<td>CMA</td>
<td>0.24</td>
<td>1.92</td>
<td>0.121</td>
<td>0.015</td>
</tr>
<tr>
<td>Momentum</td>
<td>0.61</td>
<td>4.06</td>
<td>0.047</td>
<td>0.022</td>
</tr>
<tr>
<td>Investment</td>
<td>0.29</td>
<td>1.77</td>
<td>0.099</td>
<td>0.027</td>
</tr>
<tr>
<td>Profitability</td>
<td>0.44</td>
<td>2.4</td>
<td>0.117</td>
<td>0.034</td>
</tr>
<tr>
<td>Investment Growth</td>
<td>0.7</td>
<td>1.87</td>
<td>0.102</td>
<td>0.141</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Non-traded Factors</th>
<th>Mean</th>
<th>Std</th>
<th>AR(1)</th>
<th>R-Square (Perc)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consumption growth</td>
<td>0.26</td>
<td>0.82</td>
<td>0.01</td>
<td>na</td>
</tr>
<tr>
<td>Broker-dealer Leverage</td>
<td>0.09</td>
<td>6.71</td>
<td>0.09</td>
<td>na</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Lagged instruments</th>
<th>Mean</th>
<th>Std</th>
<th>AR(1)</th>
<th>R-Square (Perc)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Old</td>
<td>Mean</td>
<td>Std</td>
<td>AR(1)</td>
<td>R-Square (Perc)</td>
</tr>
<tr>
<td>Dividend yield</td>
<td>-0.02</td>
<td>0.32</td>
<td>0.89</td>
<td>0.01</td>
</tr>
<tr>
<td>Yield spread</td>
<td>0.00</td>
<td>0.27</td>
<td>0.90</td>
<td>0.02</td>
</tr>
<tr>
<td>Term spread</td>
<td>-0.01</td>
<td>0.84</td>
<td>0.87</td>
<td>0.00</td>
</tr>
<tr>
<td>Past risk free rate</td>
<td>0.00</td>
<td>0.09</td>
<td>0.77</td>
<td>0.21</td>
</tr>
<tr>
<td>New</td>
<td>Mean</td>
<td>Std</td>
<td>AR(1)</td>
<td>R-Square (Perc)</td>
</tr>
<tr>
<td>Illiq</td>
<td>-1.75</td>
<td>0.20</td>
<td>0.84</td>
<td>0.99</td>
</tr>
<tr>
<td>New durables orders</td>
<td>0.00</td>
<td>0.04</td>
<td>0.69</td>
<td>0.28</td>
</tr>
<tr>
<td>Technical indicators</td>
<td>-0.08</td>
<td>1.47</td>
<td>0.91</td>
<td>0.55</td>
</tr>
<tr>
<td>Average Corrrelation</td>
<td>0.28</td>
<td>0.11</td>
<td>0.90</td>
<td>0.80</td>
</tr>
</tbody>
</table>
Table 2: Accuracy of Bias Adjustments for Squared Sharpe Ratios

The “true” squared Sharpe ratios are from simulations with a large number (1000*743) of time series observations. The values are stated in percent (multiplied by 100). The Average values across 5,000 simulation trials are shown for five alternative bias adjustment methods. The number of time series observations in the finite samples is 743. The JK uses the results of Jobson and Korkie (1980), and are based on a Non-central F distribution. The four adjustments for dynamic portfolios are the Chi-square, Non-central F, Odd-even and Direct Expansion. The adjustments are applied to the squared Sharpe ratios of fixed weight portfolios in Panel A and to Efficient with respect to Z portfolios in Panel B. The four lagged instruments that comprise the vector Z are described in the text. The N=25 portfolios are the 5x5 size x book/market sorts, the N=49 are industry portfolios and the N=99 combine the first two sets with 25 investment x profitability portfolios from Kenneth French.

Panel A: Fixed-weight Factor Portfolios

<table>
<thead>
<tr>
<th></th>
<th>TRUE</th>
<th>JK</th>
<th>% Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S(R_m)$</td>
<td>1.64</td>
<td>1.65</td>
<td>-1%</td>
</tr>
<tr>
<td>$S_{fix}(FF3)$</td>
<td>3.27</td>
<td>3.33</td>
<td>-2%</td>
</tr>
<tr>
<td>$S_{fix}(FF6)$</td>
<td>11.56</td>
<td>11.89</td>
<td>-3%</td>
</tr>
<tr>
<td>$S_{fix}(Q5)$</td>
<td>31.06</td>
<td>31.05</td>
<td>0%</td>
</tr>
<tr>
<td>$S_{fix}(r)$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N=25</td>
<td>16.14</td>
<td>16.50</td>
<td>-2%</td>
</tr>
<tr>
<td>N=49</td>
<td>31.87</td>
<td>32.25</td>
<td>-1%</td>
</tr>
<tr>
<td>N=99</td>
<td>76.87</td>
<td>78.35</td>
<td>-2%</td>
</tr>
</tbody>
</table>

Panel B: Efficient-with-Respect to Z Portfolios

<table>
<thead>
<tr>
<th></th>
<th>TRUE</th>
<th>No-Adj</th>
<th>Chi-Square</th>
<th>Non-Central F</th>
<th>Odd-Even</th>
<th>Direct Expansion</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_{UE}(FF3)$</td>
<td>7.13</td>
<td>8.56</td>
<td>8.10</td>
<td>7.64</td>
<td>7.58</td>
<td>7.29</td>
</tr>
<tr>
<td>$S_{UE}(FF5)$</td>
<td>15.85</td>
<td>18.61</td>
<td>17.77</td>
<td>16.93</td>
<td>16.76</td>
<td>16.20</td>
</tr>
<tr>
<td>$S_{UE}(FF6)$</td>
<td>20.04</td>
<td>23.53</td>
<td>22.47</td>
<td>21.25</td>
<td>20.99</td>
<td>20.52</td>
</tr>
<tr>
<td>$S_{UE}(Q5)$</td>
<td>33.77</td>
<td>36.72</td>
<td>35.70</td>
<td>34.35</td>
<td>34.01</td>
<td>34.12</td>
</tr>
<tr>
<td>$S_{UE}(r)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N=25</td>
<td>39.16</td>
<td>54.64</td>
<td>49.08</td>
<td>41.64</td>
<td>39.63</td>
<td>39.86</td>
</tr>
<tr>
<td>N=49</td>
<td>69.42</td>
<td>100.85</td>
<td>75.71</td>
<td>69.01</td>
<td>75.46</td>
<td>70.58</td>
</tr>
<tr>
<td>N=99</td>
<td>154.07</td>
<td>236.29</td>
<td>188.25</td>
<td>156.40</td>
<td>188.01</td>
<td>158.28</td>
</tr>
</tbody>
</table>
Table 3: Accuracy of Asymptotic Standard Deviations

A parametric bootstrap generates 1000 simulation trials, each with 743 observations. Squared Sharpe ratios and squared Sharpe ratio differences are estimated and the asymptotic standard deviations are calculated using the propositions and the Theorem I. The first columns (Empirical) are the standard deviations of the estimates across the 1,000 simulation trials. The second columns (Avg Asymptotic) are the averages of the estimated asymptotic standard deviations. Fix(r) or fix(f) refers to a mean-variance efficient portfolio that ignores the conditioning information and uses fixed weights. UE is efficient with respect to Z. The lagged instruments are the four described in the data section. The average return of a three-month Treasury bill is taken to be the zero beta rate. The N=25 portfolios are the 5x5 size x book/market sorts, the N=49 are industry portfolios and the N=99 combine the first two sets with 25 investment x profitability portfolios from Kenneth French.

### Panel A: Standard Errors for Squared Sharpe Ratio Levels

<table>
<thead>
<tr>
<th></th>
<th>Empirical (simulated)</th>
<th>Average FSW Asymptotic</th>
<th>Average BKRS Asymptotic</th>
<th>Difference FSW (% empirical)</th>
<th>Difference BKRS (% empirical)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_m$</td>
<td>0.28</td>
<td>0.28</td>
<td>0.28</td>
<td>-2%</td>
<td>-1%</td>
</tr>
<tr>
<td>$S_{fix}(FF3)$</td>
<td>0.45</td>
<td>0.40</td>
<td>0.41</td>
<td>-11%</td>
<td>-9%</td>
</tr>
<tr>
<td>$S_{fix}(FF6)$</td>
<td>0.87</td>
<td>0.80</td>
<td>0.81</td>
<td>-9%</td>
<td>-7%</td>
</tr>
<tr>
<td>$S_{UE}(FF3)$</td>
<td>0.60</td>
<td>0.58</td>
<td></td>
<td>-3%</td>
<td></td>
</tr>
<tr>
<td>$S_{UE}(FF6)$</td>
<td>1.09</td>
<td>1.08</td>
<td></td>
<td>-2%</td>
<td></td>
</tr>
<tr>
<td>$S_{fix}(r)$ N=25</td>
<td>1.02</td>
<td>0.95</td>
<td>0.95</td>
<td>-7%</td>
<td>-7%</td>
</tr>
<tr>
<td>$S_{fix}(r)$ N=49</td>
<td>1.25</td>
<td>1.24</td>
<td>1.30</td>
<td>0%</td>
<td>4%</td>
</tr>
<tr>
<td>$S_{fix}(r)$ N=99</td>
<td>2.33</td>
<td>2.48</td>
<td>2.54</td>
<td>7%</td>
<td>9%</td>
</tr>
<tr>
<td>$S_{UE}(r)$ N=25</td>
<td>1.49</td>
<td>1.38</td>
<td></td>
<td>-8%</td>
<td></td>
</tr>
<tr>
<td>$S_{UE}(r)$ N=49</td>
<td>1.91</td>
<td>1.79</td>
<td></td>
<td>-6%</td>
<td></td>
</tr>
<tr>
<td>$S_{UE}(r)$ N=99</td>
<td>3.50</td>
<td>3.33</td>
<td></td>
<td>-5%</td>
<td></td>
</tr>
</tbody>
</table>

### Panel B: Standard Errors for Squared Sharpe Ratio Differences (N=25)

<table>
<thead>
<tr>
<th></th>
<th>Empirical (simulated)</th>
<th>Average FSW Asymptotic</th>
<th>Average BKRS Asymptotic</th>
<th>Difference FSW (% empirical)</th>
<th>Difference BKRS (% empirical)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_{fix}(r) - R_m$</td>
<td>0.92</td>
<td>0.91</td>
<td>0.93</td>
<td>-1%</td>
<td>1%</td>
</tr>
<tr>
<td>$S_{fix}(r) - S_{fix}(FF3)$</td>
<td>0.93</td>
<td>0.89</td>
<td>0.91</td>
<td>-4%</td>
<td>-2%</td>
</tr>
<tr>
<td>$S_{fix}(r) - S_{fix}(FF5)$</td>
<td>0.89</td>
<td>0.89</td>
<td>0.91</td>
<td>1%</td>
<td>2%</td>
</tr>
<tr>
<td>$S_{UE}(r) - R_m$</td>
<td>1.45</td>
<td>1.35</td>
<td></td>
<td>-7%</td>
<td></td>
</tr>
<tr>
<td>$S_{UE}(r) - S_{UE}(FF3)$</td>
<td>1.46</td>
<td>1.34</td>
<td></td>
<td>-9%</td>
<td></td>
</tr>
<tr>
<td>$S_{UE}(r) - S_{UE}(FF3)$</td>
<td>1.43</td>
<td>1.34</td>
<td></td>
<td>-6%</td>
<td></td>
</tr>
<tr>
<td>$S_{UE}(r) - S_{fix}(r)$</td>
<td>1.11</td>
<td>1.02</td>
<td></td>
<td>-8%</td>
<td></td>
</tr>
</tbody>
</table>

### Panel C: Standard Errors for Squared Sharpe Ratio Differences (N=49)

...
### Panel D: Standard Errors for Squared Sharpe Ratio Differences (N=99)

<table>
<thead>
<tr>
<th></th>
<th>Empirical (simulated)</th>
<th>Average FSW Asymptotic</th>
<th>Average BKRS Asymptotic</th>
<th>Difference FSW (% empirical)</th>
<th>Difference BKRS (% empirical)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_{\text{fix}}(r) - R_m$</td>
<td>2.29</td>
<td>2.46</td>
<td>2.50</td>
<td>7%</td>
<td>9%</td>
</tr>
<tr>
<td>$S_{\text{fix}}(r) - S_{\text{fix}}(\text{FF3})$</td>
<td>2.31</td>
<td>2.45</td>
<td>2.49</td>
<td>6%</td>
<td>8%</td>
</tr>
<tr>
<td>$S_{\text{fix}}(r) - S_{\text{fix}}(\text{FF5})$</td>
<td>2.22</td>
<td>2.39</td>
<td>2.49</td>
<td>8%</td>
<td>10%</td>
</tr>
<tr>
<td>$S_{\text{UE}}(r) - R_m$</td>
<td>3.60</td>
<td>3.34</td>
<td></td>
<td>-7%</td>
<td></td>
</tr>
<tr>
<td>$S_{\text{UE}}(r) - S_{\text{UE}}(\text{FF3})$</td>
<td>3.40</td>
<td>3.29</td>
<td></td>
<td>-3%</td>
<td></td>
</tr>
<tr>
<td>$S_{\text{UE}}(r) - S_{\text{UE}}(\text{FF5})$</td>
<td>3.29</td>
<td>3.23</td>
<td></td>
<td>-2%</td>
<td></td>
</tr>
<tr>
<td>$S_{\text{UE}}(r) - S_{\text{fix}}(r)$</td>
<td>2.47</td>
<td>2.54</td>
<td></td>
<td>3%</td>
<td></td>
</tr>
</tbody>
</table>

### Panel E: Standard Errors for Squared Sharpe Ratio Differences (Factors Alone)

<table>
<thead>
<tr>
<th></th>
<th>Empirical (simulated)</th>
<th>Average FSW Asymptotic</th>
<th>Average BKRS Asymptotic</th>
<th>Difference FSW (% empirical)</th>
<th>Difference BKRS (% empirical)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_{\text{fix}}(\text{FF5}) - S_{\text{fix}}(\text{FF3})$</td>
<td>0.54</td>
<td>0.48</td>
<td>0.49</td>
<td>-12%</td>
<td>-10%</td>
</tr>
<tr>
<td>$S_{\text{fix}}(\text{FF6}) - S_{\text{fix}}(\text{FF5})$</td>
<td>0.49</td>
<td>0.47</td>
<td>0.47</td>
<td>-4%</td>
<td>-3%</td>
</tr>
<tr>
<td>$S_{\text{fix}}(\text{FF6}) - S_{\text{fix}}(\text{Q5})$</td>
<td>1.08</td>
<td>1.01</td>
<td>1.02</td>
<td>-7%</td>
<td>-6%</td>
</tr>
<tr>
<td>$S_{\text{UE}}(\text{FF5}) - S_{\text{UE}}(\text{FF3})$</td>
<td>0.65</td>
<td>0.67</td>
<td></td>
<td>2%</td>
<td></td>
</tr>
<tr>
<td>$S_{\text{UE}}(\text{FF6}) - S_{\text{UE}}(\text{FF5})$</td>
<td>0.55</td>
<td>0.55</td>
<td></td>
<td>-1%</td>
<td></td>
</tr>
<tr>
<td>$S_{\text{UE}}(\text{FF6}) - S_{\text{UE}}(\text{Q5})$</td>
<td>1.20</td>
<td>1.37</td>
<td></td>
<td>14%</td>
<td></td>
</tr>
</tbody>
</table>
Table 4: The Empirical Distributions of t-ratios

A parametric bootstrap generates 1000 simulation trials. Each set of simulated data has 743 observations. Squared Sharpe ratios, \( S(\cdot) \), and their differences are estimated, bias adjusted using the second order expansion method, and their asymptotic standard deviations are calculated using the propositions and the Theorem I. Squared t-ratios are formed as the squared adjusted Sharpe ratio or difference less its “true” value, divided by its asymptotic standard error. The true values are from simulations with 743*1000 observations. Fractiles of the empirical distribution from the 1,000 simulation trials are shown. \( \chi(1) \) are the values for a Chi distribution with one degree of freedom. Fixed weight portfolios ignore the conditioning information. UE is efficient with respect to \( Z \). The lagged instruments are the four “classical” ones. The average return of a three-month Treasury bill is taken to be the zero beta rate. The N=25 portfolios are the 5x5 size x book/market sorts, the N=49 are industry portfolios and the N=99 combine the first two sets with 25 investment x profitability portfolios from Kenneth French.

Panel A: Fractile Values of T-ratios for Squared Sharpe Ratio Levels

<table>
<thead>
<tr>
<th>Percentile: ( \chi(1) )</th>
<th>Fixed Weight Portfolios</th>
<th>Dynamic UE Portfolios</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>90%</td>
<td>95%</td>
</tr>
<tr>
<td>( S(R_m) )</td>
<td>1.65</td>
<td>1.96</td>
</tr>
<tr>
<td>( S(FF3) )</td>
<td>2.01</td>
<td>2.39</td>
</tr>
<tr>
<td>( S(FF5) )</td>
<td>1.87</td>
<td>2.18</td>
</tr>
<tr>
<td>( S(FF6) )</td>
<td>1.80</td>
<td>2.17</td>
</tr>
<tr>
<td>( S(QS5) )</td>
<td>1.67</td>
<td>2.00</td>
</tr>
<tr>
<td>( S(r) ) ( N=25 )</td>
<td>1.73</td>
<td>2.08</td>
</tr>
<tr>
<td>( S(r) ) ( N=49 )</td>
<td>1.57</td>
<td>1.84</td>
</tr>
<tr>
<td>( S(r) ) ( N=99 )</td>
<td>1.59</td>
<td>1.96</td>
</tr>
</tbody>
</table>

Panel B: Squared Sharpe Ratio Differences (N=25)

<table>
<thead>
<tr>
<th>Percentile: ( \chi(1) )</th>
<th>Fixed Weight Portfolios</th>
<th>Dynamic UE Portfolios</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>90%</td>
<td>95%</td>
</tr>
<tr>
<td>( \chi(1) )</td>
<td>1.65</td>
<td>1.96</td>
</tr>
<tr>
<td>( S(r) - R_m )</td>
<td>1.62</td>
<td>1.96</td>
</tr>
<tr>
<td>( S(r) - S(FF3) )</td>
<td>1.76</td>
<td>2.05</td>
</tr>
<tr>
<td>( S(r) - S(FF5) )</td>
<td>1.63</td>
<td>1.99</td>
</tr>
<tr>
<td>( S_{UE}(r) - S_{fix}(r) )</td>
<td>1.83</td>
<td>2.13</td>
</tr>
</tbody>
</table>

| \( S_{UE}(r) - S_{fix}(r) \) | NA  | NA  | NA  |
### Panel C: Squared Sharpe Ratio Differences (N=49)

<table>
<thead>
<tr>
<th></th>
<th>Fixed Weight Portfolios</th>
<th>Dynamic UE Portfolios</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentile:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \chi(1) )</td>
<td>90%  95%  98%</td>
<td>90%  95%  98%</td>
</tr>
<tr>
<td></td>
<td>1.65  1.96  2.33</td>
<td>1.65  1.96  2.33</td>
</tr>
<tr>
<td>( S(r) - Rm )</td>
<td>1.96  2.25  2.57</td>
<td>1.81  2.01  2.32</td>
</tr>
<tr>
<td>( S(r) - S(\text{FF3}) )</td>
<td>1.71  2.00  2.34</td>
<td>1.71  2.05  2.28</td>
</tr>
<tr>
<td>( S(r) - S(\text{FF5}) )</td>
<td>1.65  1.98  2.38</td>
<td>1.69  1.95  2.29</td>
</tr>
<tr>
<td>( S_{\text{UE}}(r) - S_{\text{fix}}(r) )</td>
<td>1.83  2.13  2.47</td>
<td>NA  NA  NA</td>
</tr>
</tbody>
</table>

### Panel D: Squared Sharpe Ratio Differences (N=99)

<table>
<thead>
<tr>
<th></th>
<th>Fixed Weight Portfolios</th>
<th>Dynamic UE Portfolios</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentile:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \chi(1) )</td>
<td>90%  95%  98%</td>
<td>90%  95%  98%</td>
</tr>
<tr>
<td></td>
<td>1.65  1.96  2.33</td>
<td>1.65  1.96  2.33</td>
</tr>
<tr>
<td>( S(r) - Rm )</td>
<td>1.54  1.84  2.23</td>
<td>2.01  2.41  2.73</td>
</tr>
<tr>
<td>( S(r) - S(\text{FF3}) )</td>
<td>1.59  1.88  2.25</td>
<td>1.83  2.13  2.62</td>
</tr>
<tr>
<td>( S(r) - S(\text{FF5}) )</td>
<td>1.57  1.91  2.25</td>
<td>1.79  2.13  2.56</td>
</tr>
<tr>
<td>( S_{\text{UE}}(r) - S_{\text{fix}}(r) )</td>
<td>1.80  2.07  2.44</td>
<td>NA  NA  NA</td>
</tr>
</tbody>
</table>

### Panel E: Squared Sharpe Ratio Differences for Factors Alone

<table>
<thead>
<tr>
<th></th>
<th>Fixed Weight Portfolios</th>
<th>Dynamic UE Portfolios</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentile:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \chi(1) )</td>
<td>90%  95%  98%</td>
<td>90%  95%  98%</td>
</tr>
<tr>
<td></td>
<td>1.65  1.96  2.33</td>
<td>1.65  1.96  2.33</td>
</tr>
<tr>
<td>( S(\text{FF5}) - S(\text{FF3}) )</td>
<td>1.90  2.32  3.02</td>
<td>1.60  1.96  2.53</td>
</tr>
<tr>
<td>( S(\text{FF6}) - S(\text{FF5}) )</td>
<td>1.88  2.43  3.42</td>
<td>1.79  2.27  2.88</td>
</tr>
<tr>
<td>( S(\text{FF6}) - S(Q5) )</td>
<td>1.74  2.00  2.55</td>
<td>1.41  1.75  2.03</td>
</tr>
</tbody>
</table>
Table 5: Squared Sharpe Ratios for Mimicking Portfolios

A parametric bootstrap generates 1000 simulation trials. Each set of simulated data has $T=743$ observations for consumption growth and $T=587$ for the leverage risk factor. The squared Sharpe ratios of mimicking portfolios are estimated and shown, Unadjusted and Adjusted using the JK (1980) bias adjustment. The values are stated in percent (multiplied by 100). The true values are from simulations with $T=1,000$ observations. Their asymptotic standard deviations are calculated using the propositions and the Theorem I. The Average Asymptotic value is taken across the 1,000 simulation trials. The empirical standard error is the standard deviation of the Sharpe ratio point estimates taken across the 1,000 simulation trials. The lagged instruments are the four classical instruments described in the data section. The average return of a three-month Treasury bill is taken to be the zero beta rate. The mimicking portolios are formed from 25 size x value portfolios, with either fixed weights (fix) or efficient portfolio weights with respect to the lagged information (UE).

<table>
<thead>
<tr>
<th></th>
<th>Squared Sharpe Ratio Levels</th>
<th>Standard Errors</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TRUE</td>
<td>No-Adj</td>
<td>Adj</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consumption (fixed)</td>
<td>1.41</td>
<td>1.34</td>
<td>1.20</td>
</tr>
<tr>
<td>Consumption (UE)</td>
<td>1.40</td>
<td>1.65</td>
<td>1.51</td>
</tr>
<tr>
<td>Leverage (fixed)</td>
<td>2.80</td>
<td>3.63</td>
<td>3.44</td>
</tr>
<tr>
<td>Leverage (UE)</td>
<td>2.77</td>
<td>3.67</td>
<td>3.48</td>
</tr>
</tbody>
</table>
**Table 6: Relative Tests of Factor Models**

The test statistic is the difference in the bias-adjusted squared Sharpe ratios (not multiplied by 100) for the test assets and the factors versus the factors alone. The t-ratios are in parentheses. The factor model abbreviations and test asset portfolios are described in the text. Monthly Sharpe ratios are computed using the average Treasury bill return of 0.39 percent per month as the zero-beta rate. The dynamic models trade optimally using the four classical lagged instruments in Panel B and using the modern instruments previously described in Panel C. The sample period is January 1967 to December, 2020.

### Panel A: Fixed weight Models

<table>
<thead>
<tr>
<th>Portfolio Type</th>
<th>CAPM</th>
<th>FF3</th>
<th>FF5</th>
<th>FF6</th>
<th>Q4</th>
<th>Q5</th>
</tr>
</thead>
<tbody>
<tr>
<td>25 size x value portfolios:</td>
<td>0.16</td>
<td>0.13</td>
<td>0.11</td>
<td>0.09</td>
<td>0.12</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>(4.03)</td>
<td>(3.46)</td>
<td>(2.91)</td>
<td>(2.58)</td>
<td>(3.13)</td>
<td>(2.73)</td>
</tr>
<tr>
<td>25 Investment x productivity:</td>
<td>0.06</td>
<td>0.05</td>
<td>-0.01</td>
<td>-0.01</td>
<td>0.00</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>(2.35)</td>
<td>(1.89)</td>
<td>(-0.27)</td>
<td>(-0.72)</td>
<td>(0.20)</td>
<td>(0.27)</td>
</tr>
<tr>
<td>32 size x value portfolios x prod:</td>
<td>0.09</td>
<td>0.07</td>
<td>0.05</td>
<td>0.04</td>
<td>0.05</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>(2.73)</td>
<td>(2.23)</td>
<td>(1.49)</td>
<td>(1.42)</td>
<td>(1.69)</td>
<td>(1.56)</td>
</tr>
<tr>
<td>49 Industry:</td>
<td>0.08</td>
<td>0.12</td>
<td>0.17</td>
<td>0.15</td>
<td>0.14</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>(2.59)</td>
<td>(3.08)</td>
<td>(3.84)</td>
<td>(3.44)</td>
<td>(2.98)</td>
<td>(2.55)</td>
</tr>
</tbody>
</table>

### Panel B: Dynamic Models using Classical Instruments

<table>
<thead>
<tr>
<th>Portfolio Type</th>
<th>CAPM</th>
<th>FF3</th>
<th>FF5</th>
<th>FF6</th>
<th>Q4</th>
<th>Q5</th>
</tr>
</thead>
<tbody>
<tr>
<td>25 size x value portfolios:</td>
<td>0.24</td>
<td>0.19</td>
<td>0.17</td>
<td>0.15</td>
<td>0.20</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>(4.68)</td>
<td>(3.94)</td>
<td>(3.58)</td>
<td>(3.29)</td>
<td>(3.86)</td>
<td>(3.22)</td>
</tr>
<tr>
<td>25 Investment x productivity:</td>
<td>0.11</td>
<td>0.10</td>
<td>0.03</td>
<td>0.02</td>
<td>0.06</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>(2.95)</td>
<td>(2.53)</td>
<td>(0.84)</td>
<td>(0.64)</td>
<td>(1.72)</td>
<td>(1.58)</td>
</tr>
<tr>
<td>32 size x value portfolios x prod:</td>
<td>0.18</td>
<td>0.15</td>
<td>0.11</td>
<td>0.11</td>
<td>0.16</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td>(3.76)</td>
<td>(3.03)</td>
<td>(2.53)</td>
<td>(2.46)</td>
<td>(3.07)</td>
<td>(2.70)</td>
</tr>
<tr>
<td>49 Industry:</td>
<td>0.20</td>
<td>0.21</td>
<td>0.26</td>
<td>0.24</td>
<td>0.26</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>(3.89)</td>
<td>(3.68)</td>
<td>(4.19)</td>
<td>(3.93)</td>
<td>(4.03)</td>
<td>(3.71)</td>
</tr>
</tbody>
</table>

### Panel C: Dynamic Models using Modern Instruments

<table>
<thead>
<tr>
<th>Portfolio Type</th>
<th>CAPM</th>
<th>FF3</th>
<th>FF5</th>
<th>FF6</th>
<th>Q4</th>
<th>Q5</th>
</tr>
</thead>
<tbody>
<tr>
<td>25 size x value portfolios:</td>
<td>0.32</td>
<td>0.27</td>
<td>0.24</td>
<td>0.24</td>
<td>0.28</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td>(5.62)</td>
<td>(4.94)</td>
<td>(4.45)</td>
<td>(4.42)</td>
<td>(4.78)</td>
<td>(4.11)</td>
</tr>
<tr>
<td>25 Investment x productivity:</td>
<td>0.23</td>
<td>0.21</td>
<td>0.15</td>
<td>0.14</td>
<td>0.18</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>(4.80)</td>
<td>(4.59)</td>
<td>(3.42)</td>
<td>(3.35)</td>
<td>(3.97)</td>
<td>(3.29)</td>
</tr>
<tr>
<td>32 size x value portfolios x prod:</td>
<td>0.21</td>
<td>0.18</td>
<td>0.15</td>
<td>0.15</td>
<td>0.19</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>(3.94)</td>
<td>(3.35)</td>
<td>(2.83)</td>
<td>(2.89)</td>
<td>(3.30)</td>
<td>(2.96)</td>
</tr>
<tr>
<td>49 Industry:</td>
<td>0.47</td>
<td>0.49</td>
<td>0.53</td>
<td>0.53</td>
<td>0.52</td>
<td>0.50</td>
</tr>
</tbody>
</table>
Table 7: Relative Tests of Factor Models with an Estimated Zero Beta Rate

The test statistic is the difference in bias-adjusted squared Sharpe ratios for the test assets and the factors versus the factors alone. The factor model abbreviations and test asset portfolios are described in the text. Monthly Sharpe ratios are computed using the estimated zero beta rate, assuming no risk-free asset exists, and shown as percent per month. The dynamic models trade optimally using the four lagged instruments previously described. The sample period is January 1967 to December, 2020.

Panel A: Fixed weight Models

<table>
<thead>
<tr>
<th></th>
<th>CAPM</th>
<th>FF3</th>
<th>FF5</th>
<th>FF6</th>
<th>Q4</th>
<th>Q5</th>
</tr>
</thead>
<tbody>
<tr>
<td>25 size x value portfolios:</td>
<td>0.05</td>
<td>0.07</td>
<td>0.06</td>
<td>0.05</td>
<td>0.06</td>
<td>0.08</td>
</tr>
<tr>
<td>Zero Beta rate</td>
<td>0.0155</td>
<td>0.0038</td>
<td>0.0036</td>
<td>0.0035</td>
<td>0.0028</td>
<td>0.0023</td>
</tr>
<tr>
<td>25 Investment x productivity:</td>
<td>0.03</td>
<td>0.05</td>
<td>-0.01</td>
<td>-0.02</td>
<td>0.00</td>
<td>-0.01</td>
</tr>
<tr>
<td>Zero Beta rate</td>
<td>0.0152</td>
<td>0.0052</td>
<td>0.0038</td>
<td>0.0036</td>
<td>0.0043</td>
<td>0.0042</td>
</tr>
<tr>
<td>32 size x value portfolios x prod:</td>
<td>0.03</td>
<td>0.04</td>
<td>0.02</td>
<td>0.02</td>
<td>0.03</td>
<td>0.04</td>
</tr>
<tr>
<td>Zero Beta rate</td>
<td>0.0126</td>
<td>0.0042</td>
<td>0.0039</td>
<td>0.0037</td>
<td>0.0036</td>
<td>0.0033</td>
</tr>
<tr>
<td>49 Industry:</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.09</td>
<td>0.08</td>
<td>0.08</td>
</tr>
<tr>
<td>Zero Beta rate</td>
<td>0.0083</td>
<td>0.0070</td>
<td>0.0064</td>
<td>0.0060</td>
<td>0.0076</td>
<td>0.0067</td>
</tr>
</tbody>
</table>

Panel B: Dynamic Models using Classical Instruments

<table>
<thead>
<tr>
<th></th>
<th>CAPM</th>
<th>FF3</th>
<th>FF5</th>
<th>FF6</th>
<th>Q4</th>
<th>Q5</th>
</tr>
</thead>
<tbody>
<tr>
<td>25 size x value portfolios:</td>
<td>0.11</td>
<td>0.11</td>
<td>0.11</td>
<td>0.09</td>
<td>0.14</td>
<td>0.15</td>
</tr>
<tr>
<td>Zero Beta rate</td>
<td>0.0155</td>
<td>0.0039</td>
<td>0.0037</td>
<td>0.0036</td>
<td>0.0028</td>
<td>0.0023</td>
</tr>
<tr>
<td>25 Investment x productivity:</td>
<td>0.03</td>
<td>0.05</td>
<td>-0.02</td>
<td>-0.03</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>Zero Beta rate</td>
<td>0.0154</td>
<td>0.0052</td>
<td>0.0038</td>
<td>0.0036</td>
<td>0.0042</td>
<td>0.0042</td>
</tr>
<tr>
<td>32 size x value portfolios x prod:</td>
<td>0.10</td>
<td>0.09</td>
<td>0.06</td>
<td>0.06</td>
<td>0.11</td>
<td>0.12</td>
</tr>
<tr>
<td>Zero Beta rate</td>
<td>0.0128</td>
<td>0.0042</td>
<td>0.0039</td>
<td>0.0038</td>
<td>0.0036</td>
<td>0.0033</td>
</tr>
<tr>
<td>49 Industry:</td>
<td>0.19</td>
<td>0.16</td>
<td>0.15</td>
<td>0.17</td>
<td>0.19</td>
<td>0.20</td>
</tr>
<tr>
<td>Zero Beta rate</td>
<td>0.0084</td>
<td>0.0071</td>
<td>0.0065</td>
<td>0.0060</td>
<td>0.0076</td>
<td>0.0067</td>
</tr>
</tbody>
</table>
### Panel C: Dynamic Models using Modern Instruments

<table>
<thead>
<tr>
<th></th>
<th>CAPM</th>
<th>FF3</th>
<th>FF5</th>
<th>FF6</th>
<th>Q4</th>
<th>Q5</th>
</tr>
</thead>
<tbody>
<tr>
<td>25 size x value portfolios:</td>
<td>0.12</td>
<td>0.12</td>
<td>0.10</td>
<td>0.09</td>
<td>0.14</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>(2.83)</td>
<td>(1.65)</td>
<td>(1.03)</td>
<td>(0.85)</td>
<td>(2.01)</td>
<td>(1.57)</td>
</tr>
<tr>
<td>Zero Beta rate</td>
<td>0.0162</td>
<td>0.0039</td>
<td>0.0037</td>
<td>0.0036</td>
<td>0.0029</td>
<td>0.0025</td>
</tr>
<tr>
<td>25 Investment x productivity:</td>
<td>0.07</td>
<td>0.09</td>
<td>0.02</td>
<td>0.01</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>(1.89)</td>
<td>(2.20)</td>
<td>(0.40)</td>
<td>(0.23)</td>
<td>(0.75)</td>
<td>(0.47)</td>
</tr>
<tr>
<td>Zero Beta rate</td>
<td>0.0154</td>
<td>0.0053</td>
<td>0.0038</td>
<td>0.0037</td>
<td>0.0044</td>
<td>0.0044</td>
</tr>
<tr>
<td>32 size x value portfolios x prod:</td>
<td>0.08</td>
<td>0.07</td>
<td>0.05</td>
<td>0.04</td>
<td>0.10</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>(1.79)</td>
<td>(1.13)</td>
<td>(0.49)</td>
<td>(0.38)</td>
<td>(1.57)</td>
<td>(1.25)</td>
</tr>
<tr>
<td>Zero Beta rate</td>
<td>0.0131</td>
<td>0.0042</td>
<td>0.0039</td>
<td>0.0038</td>
<td>0.0037</td>
<td>0.0034</td>
</tr>
<tr>
<td>49 Industry:</td>
<td>0.38</td>
<td>0.37</td>
<td>0.36</td>
<td>0.36</td>
<td>0.37</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td>(5.51)</td>
<td>(5.72)</td>
<td>(5.62)</td>
<td>(5.56)</td>
<td>(6.16)</td>
<td>(6.36)</td>
</tr>
<tr>
<td>Zero Beta rate</td>
<td>0.0088</td>
<td>0.0072</td>
<td>0.0065</td>
<td>0.0061</td>
<td>0.0080</td>
<td>0.0073</td>
</tr>
</tbody>
</table>
Table 8: Direct Factor Model Comparisons

The test statistic is the difference in bias-adjusted squared Sharpe ratios (not multiplied by 100) for the first model less the second model. The factors are held with fixed weights over time (no instruments) or dynamically traded using either the four classical or the four modern instruments. The t-ratios are the differences divided by the asymptotic standard errors for the difference. The factor model abbreviations and test asset portfolios are described in the text. FF6* replaces the HML factor in FF6 factors version. The sample period is January 1967 to December, 2020.

<table>
<thead>
<tr>
<th></th>
<th>No Instruments</th>
<th>Classical Instruments</th>
<th>Modern Instruments</th>
</tr>
</thead>
<tbody>
<tr>
<td>FF3 – R_m</td>
<td>0.03</td>
<td>0.05</td>
<td>0.06</td>
</tr>
<tr>
<td>(t-ratio)</td>
<td>(1.73)</td>
<td>(2.45)</td>
<td>(2.74)</td>
</tr>
<tr>
<td>FF5 – FF3</td>
<td>0.06</td>
<td>0.08</td>
<td>0.08</td>
</tr>
<tr>
<td>(t-ratio)</td>
<td>(2.38)</td>
<td>(2.74)</td>
<td>(2.44)</td>
</tr>
<tr>
<td>FF6 – FF5</td>
<td>0.03</td>
<td>0.03</td>
<td>0.06</td>
</tr>
<tr>
<td>(t-ratio)</td>
<td>(1.37)</td>
<td>(1.46)</td>
<td>(2.40)</td>
</tr>
<tr>
<td>Q5 – Q4</td>
<td>0.20</td>
<td>0.20</td>
<td>0.21</td>
</tr>
<tr>
<td>(t-ratio)</td>
<td>(3.75)</td>
<td>(3.80)</td>
<td>(3.81)</td>
</tr>
<tr>
<td>Q4 – FF5</td>
<td>0.08</td>
<td>0.03</td>
<td>0.06</td>
</tr>
<tr>
<td>(t-ratio)</td>
<td>(2.20)</td>
<td>(0.81)</td>
<td>(1.36)</td>
</tr>
<tr>
<td>Q5 – FF6</td>
<td>0.25</td>
<td>0.21</td>
<td>0.20</td>
</tr>
<tr>
<td>(t-ratio)</td>
<td>(4.05)</td>
<td>(3.23)</td>
<td>(3.09)</td>
</tr>
<tr>
<td>Q5 – FF6*</td>
<td>0.22</td>
<td>0.19</td>
<td>0.19</td>
</tr>
<tr>
<td>(t-ratio)</td>
<td>(3.32)</td>
<td>(2.86)</td>
<td>(2.81)</td>
</tr>
<tr>
<td>Q5 – M4</td>
<td>0.16</td>
<td>0.13</td>
<td>0.15</td>
</tr>
<tr>
<td>(t-ratio)</td>
<td>(2.66)</td>
<td>(2.16)</td>
<td>(2.32)</td>
</tr>
</tbody>
</table>
Table 5: Squared Sharpe Ratios for Mimicking Portfolios

A parametric bootstrap generates 1000 simulation trials. Each set of simulated data has $T=743$ observations for consumption growth and $T=587$ for the leverage risk factor and $T=743$ for the Chen, Roll and Ross (1986) factors. The squared Sharpe ratios of mimicking portfolios are estimated and shown, Unadjusted and Adjusted using the JK (1980) bias adjustment. The values are stated in percent (multiplied by 100). The true values are from simulations with $T*1,000$ observations. Asymptotic standard deviations are calculated using the propositions and Theorem I. The Average Asymptotic value is the average taken across the 1,000 simulation trials. The empirical standard error is the standard deviation of the Sharpe ratio bias-adjusted point estimates taken across the 1,000 simulation trials. The lagged instruments are described in the data section. The average return of a three-month Treasury bill is taken to be the zero beta rate. The mimicking portfolios are formed from 25 size x value portfolios, with either fixed weights (fixed) or time-varying weights (Z). The bottom rows use fixed-weight mimicking portfolios for the CRR factors, combined with time-varying weights (with Z).

<table>
<thead>
<tr>
<th></th>
<th>Squared Sharpe Ratio Levels</th>
<th>Standard Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TRUE</td>
<td>No-Adj</td>
</tr>
<tr>
<td>Consumption (fixed)</td>
<td>1.54</td>
<td>1.35</td>
</tr>
<tr>
<td>Consumption (with Z)</td>
<td>1.60</td>
<td>1.69</td>
</tr>
<tr>
<td>Leverage (fixed)</td>
<td>2.47</td>
<td>2.70</td>
</tr>
<tr>
<td>Leverage (with Z)</td>
<td>2.43</td>
<td>2.73</td>
</tr>
<tr>
<td>Change Expected Inflation (fixed)</td>
<td>0.28</td>
<td>0.58</td>
</tr>
<tr>
<td>Change Expected Inflation (with Z)</td>
<td>0.27</td>
<td>1.12</td>
</tr>
<tr>
<td>Industry Production (fixed)</td>
<td>2.05</td>
<td>1.86</td>
</tr>
<tr>
<td>Industry Production (with Z)</td>
<td>1.98</td>
<td>2.09</td>
</tr>
<tr>
<td>Unexpected Inflation (fixed)</td>
<td>0.03</td>
<td>0.37</td>
</tr>
<tr>
<td>Unexpected Inflation (with Z)</td>
<td>0.04</td>
<td>0.80</td>
</tr>
<tr>
<td>Three CRR factors (fixed)</td>
<td>3.03</td>
<td>3.02</td>
</tr>
<tr>
<td>Three CRR factors (with Z)</td>
<td>4.61</td>
<td>6.13</td>
</tr>
</tbody>
</table>