Supplementary Material to

Optimal Portfolio Choice with Fat Tails and Parameter Uncertainty

Raymond Kan and Nathan Lassance

This supplementary material to the main paper contains four sections. Section A illustrates the estimation accuracy of the number of degrees of freedom ν and the distribution of τ_t. Section B studies the impact of fat tails on the accuracy of the adjusted estimators of θ^2 and ψ^2. Section C reports the tables containing the results for the additional empirical tests. Section D contains the proofs of all theoretical results in the main body of the paper.

A. Accuracy of Estimators of ν and τ_t

In this section, we illustrate the estimation accuracy of the number of degrees of freedom ν and the distribution of τ_t, which underlie the two calibration methods proposed in Section V to estimate the optimal two-fund and three-fund combination coefficients.

Figure A.1 studies the estimation accuracy for ν, which we estimate by maximum likelihood. We set $N = 25$, a population value of $\nu = (4,6,8)$, and we depict boxplots of $\hat{\nu}$ across 10,000 simulations of multivariate t-distributed returns for a sample size $T = (60,120,240)$. We set $(\mu,\Sigma) = (0_N, I_N)$ without loss of generality because ν does not depend on (μ,Σ). Figure A.1 shows that as ν increases, and thus the returns are closer to multivariate normal, it becomes more difficult to estimate ν. Specifically, the boxplots get wider as ν increases. However, in comparison to the volatility of $\hat{\nu}$, the bias of $\hat{\nu}$ is more reasonable, and close to zero for $T = 120$ and 240.

Figure A.2 illustrates how the sample distribution of τ_t by El Karoui (2010, 2013), $\hat{\tau}_t$ in (60), converges to the true distribution of τ_t as N increases. We assume returns are multivariate t-distributed, in which case we can show that the exact density function of τ_t in (2) is

$$f_{\tau_t}(x) = \frac{(\nu/2-1)^{\frac{\nu}{2}} e^{-\frac{\nu-2}{2x}}}{\Gamma(\nu/2) x^{\frac{\nu+2}{2}}}. \tag{A1}$$

Figure A.1: Boxplots of Estimates of the Number of Degrees of Freedom ν.

This figure depicts boxplots of maximum-likelihood estimates of the number of degrees of freedom ν of the multivariate t-distribution. The boxplots are obtained by simulating 10,000 times T return vectors from a multivariate t-distribution with $(\mu,\Sigma) = (0_N, I_N)$, $N = 25$, $\nu = (4,6,8)$, and $T = (60,120,240)$. The dotted horizontal lines depict the true value of ν.

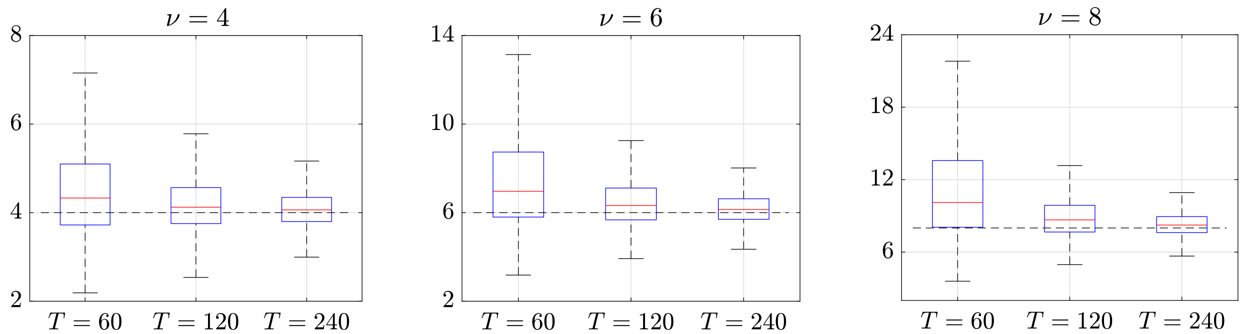
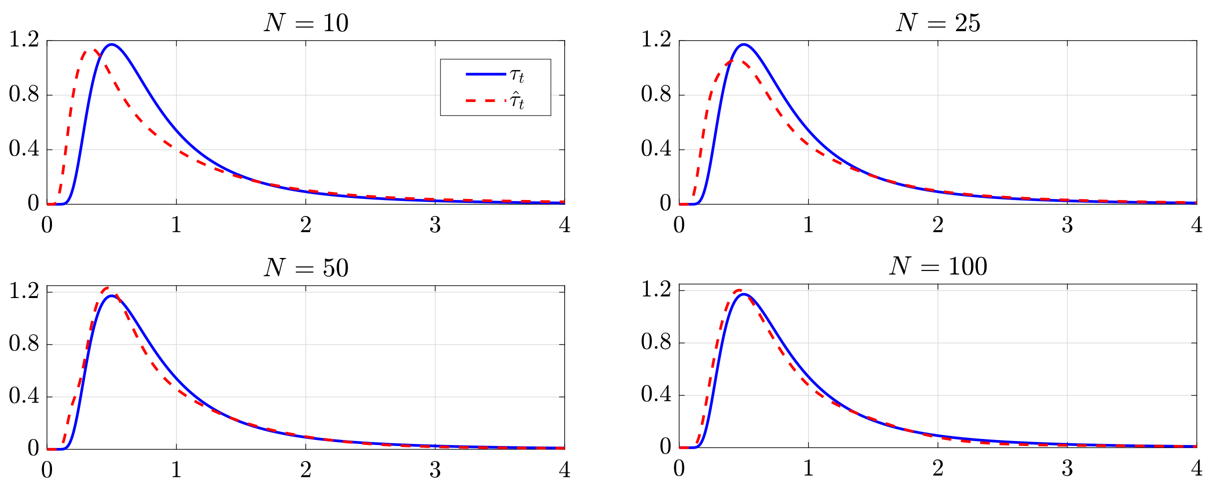


Figure A.2: Comparison of the True Distribution of τ_t with the Sample Distribution.

This figure compares the sample distribution of τ_t by El Karoui (2010, 2013), $\hat{\tau}_t$ in (60), with the true distribution of τ_t. We assume returns are multivariate t-distributed, in which case the exact density function of τ_t is given by (A1). The density function of $\hat{\tau}_t$ is found using a kernel density estimator. We set $T = 120$, $\nu = 6$, and an increasing number of assets N that goes from 10 to 100.



Then, we set $T = 120$, $\nu = 6$, and we compare the true density function of τ_t in (A1) with that of $\hat{\tau}_t$ found using a kernel density estimator. We do the comparison for an increasing number of assets N that goes from 10 to 100. Figure A.2 shows indeed that the sample distribution and the true distribution get closer to one another as N increases even for a finite T. Moreover, the sample distribution is reasonably accurate even for rather small values of N such as $N = 25$.

B. Impact of Fat Tails on Adjusted Estimators of θ^2 and ψ^2

In the main body of the paper, we estimate θ^2 and ψ^2 via their adjusted estimators θ_a^2 and ψ_a^2 in (54)–(55). These estimators are proposed by Kan and Zhou (2007) and are designed to have minimum root mean square error (RMSE). Unlike the unbiased estimators $\hat{\theta}_{unb}^2$ and $\hat{\psi}_{unb}^2$, given by the first term in (54)–(55), the adjusted estimators are non-negative. Moreover, the adjusted estimators deliver a lower RMSE than the trimmed estimators $\max(\hat{\theta}_{unb}^2, 0)$ and $\max(\hat{\psi}_{unb}^2, 0)$.

However, the adjusted estimators are derived under the multivariate normal distributional assumption, whereas we assume that returns are multivariate elliptical. Therefore, it is of interest to study how fat tails impact the RMSE of $\hat{\theta}_a^2$ and $\hat{\psi}_a^2$. For that purpose, we conduct the following simulation. We simulate $M = 1{,}000{,}000$ times T returns from a multivariate t-distribution with ν degrees of freedom and (μ, Σ) calibrated to a dataset of $N = 25$ portfolios of firms sorted on size and book-to-market spanning July 1926 to July 2023. This choice of (μ, Σ) yields $\theta = 0.302$ and $\psi = 0.250$ in the population. For each simulation $m = 1, \ldots, M$, we obtain adjusted estimates $\hat{\theta}_{a,m}^2$ and $\hat{\psi}_{a,m}^2$, and compute the RMSE as

$$\text{(A2)} \qquad \text{RMSE}(\hat{\theta}_a^2) = \sqrt{\frac{1}{M}\sum_{m=1}^{M}(\hat{\theta}_{a,m}^2 - \theta^2)^2} \quad \text{and} \quad \text{RMSE}(\hat{\psi}_a^2) = \sqrt{\frac{1}{M}\sum_{m=1}^{M}(\hat{\psi}_{a,m}^2 - \psi^2)^2}.$$

Figure A.3: Root Mean Squared Error of θ_a^2 and ψ_a^2

This figure depicts the root mean squared error (RMSE) of the adjusted estimators $\hat{\theta}_a^2$ and $\hat{\psi}_a^2$ when the asset returns are multivariate t-distributed with ν degrees of freedom, where ν varies between 4 and 20. We consider a sample size $T = 60$, 120, and 240 months. We calibrate the population value of θ and ψ to a dataset of $N = 25$ portfolios of firms sorted on size and book-to-market spanning July 1926 to July 2023, which yields $\theta = 0.302$ and $\psi = 0.250$. The RMSE is obtained over one million simulations.

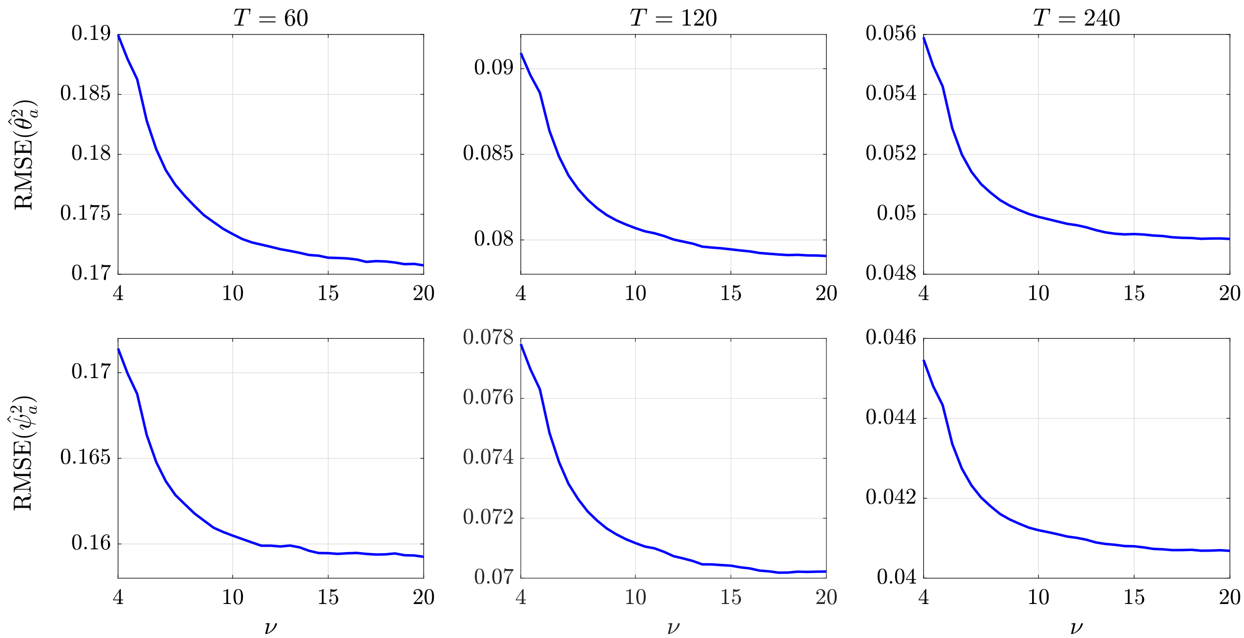


Figure A.3 depicts RMSE($\hat{\theta}_a^2$) and RMSE($\hat{\psi}_a^2$) for ν varying between 4 and 20 and $T = 60$, 120, and 240 months. Figure A.3 shows that the RMSE does not increase much as ν gets smaller and tails get fatter. For example, when going from $\nu = 20$ (close to normal) to $\nu = 6$ (excess kurtosis of three), RMSE($\hat{\theta}_a^2$) goes from 0.171 to 0.180 for $T = 60$, 0.078 to 0.084 for $T = 120$, and 0.049 to 0.052 for $T = 240$, respectively. The conclusion is similar for RMSE($\hat{\psi}_a^2$). Moreover, Figure A.3 shows that the RMSE of $\hat{\theta}_a^2$ and $\hat{\psi}_a^2$ is particularly large when $T = 60$ months, which worsens the estimation accuracy for the two-fund and three-fund combination coefficients. This partly explains why, in Figure 3 of the main body of the manuscript, the empirical performance is generally much worse (and sometimes negative) when $T = 60$ months relative to $T = 120$ and 240 months.

C. Tables for the Additional Empirical Tests

In this section, we report tables containing additional empirical results. Specifically, Table C.1 reports the skewness and excess kurtosis of the two-fund and three-fund rules, Table C.2 reports the in-sample versus out-of-sample performance of the two-fund and three-fund rules discussed in Section VI.C.1, Table C.3 reports the results for the combination of the sample GMV portfolio with the risk-free asset discussed in Section VI.C.2, and Table C.4 reports the results for daily data discussed in Section VI.C.3.

Table C.1: Skewness and Excess Kurtosis of Two-Fund and Three-Fund Rules.

This table reports the monthly skewness and excess kurtosis of the net-of-cost out-of-sample returns of the two-fund and three-fund rules across the six datasets described in Section VI.B. The combination coefficients are calibrated either to the multivariate normal distribution or to the multivariate elliptical distribution using the exact finite-sample formula. Formulas for the estimated two-fund and three-fund combination coefficients are available in Table 1. See the notes of Table 3 in the main body of the paper for details.

		10MOM		**16ANOM**		**25SBETA**	
		Normal	Elliptical	Normal	Elliptical	Normal	Elliptical
Two-fund rule							
$T = 60$	Skewness	1.675	0.722	0.263	0.281	−0.912	−0.832
	Exc. kurtosis	18.60	9.977	4.266	4.774	16.98	12.40
$T = 120$	Skewness	0.727	0.662	0.381	0.460	−0.285	−0.509
	Exc. kurtosis	5.334	4.554	3.224	3.575	12.62	8.634
$T = 240$	Skewness	−0.153	−0.171	0.829	0.794	−0.914	−0.975
	Exc. kurtosis	4.179	3.952	3.310	3.088	4.483	4.742
Three-fund rule							
$T = 60$	Skewness	1.255	0.635	0.115	0.181	−0.126	−0.692
	Exc. kurtosis	12.74	7.936	4.508	5.508	13.28	10.289
$T = 120$	Skewness	0.472	0.480	−0.259	−0.172	−0.448	−0.716
	Exc. kurtosis	3.558	3.465	3.860	3.904	10.26	7.234
$T = 240$	Skewness	−0.237	−0.240	0.307	0.384	−1.046	−1.102
	Exc. kurtosis	3.298	3.391	2.089	2.134	3.502	4.122
		25SBTM		**25OPINV**		**30IND**	
		Normal	Elliptical	Normal	Elliptical	Normal	Elliptical
Two-fund rule							
$T = 60$	Skewness	0.786	1.101	−0.990	−0.666	−1.547	0.185
	Exc. kurtosis	16.90	15.61	17.08	15.06	49.79	25.36
$T = 120$	Skewness	0.833	0.691	−0.560	−0.298	0.756	0.730
	Exc. kurtosis	5.131	4.794	7.867	7.259	7.602	7.398
$T = 240$	Skewness	0.433	0.393	−0.868	−1.001	−0.716	−0.704
	Exc. kurtosis	7.349	6.358	10.62	10.12	7.515	7.364
Three-fund rule							
$T = 60$	Skewness	0.157	0.945	−0.811	−0.639	−0.966	0.252
	Exc. kurtosis	15.54	15.32	12.12	14.20	22.06	16.17
$T = 120$	Skewness	0.461	0.386	0.079	0.015	0.733	0.926
	Exc. kurtosis	3.342	3.261	3.703	3.416	8.738	10.25
$T = 240$	Skewness	0.368	0.345	−0.408	−0.601	−0.151	−0.226
	Exc. kurtosis	4.445	4.656	6.241	6.555	4.297	5.007

Table C.2: In-Sample versus Out-of-Sample Performance of Two-Fund and Three-Fund Rules.

This table reports, for the two-fund and three-fund rules and across the six datasets described in Section VI.B, the difference between 1) the average in-sample annualized mean return, volatility, and utility over all estimation windows of size $T = 60$, 120, and 240 months versus 2) the corresponding out-of-sample realized statistic. The difference for the volatility is always negative and we report it positively. Thus, the lower the differences, the better. The combination coefficients are calibrated either to the multivariate normal, t, or elliptical distribution using the exact finite-sample formula. Formulas for the estimated two-fund and three-fund combination coefficients are available in Table 1. See the notes of Table 3 in the main body of the paper for details.

		10MOM			**16ANOM**			**25SBETA**		
		Normal	t	Elliptical	Normal	t	Elliptical	Normal	t	Elliptical
Two-fund rule										
$T = 60$	Mean	0.757	0.704	0.577	1.483	1.391	0.961	1.152	1.124	0.672
	Vol	0.567	0.523	0.352	0.609	0.581	0.388	0.716	0.704	0.391
	Utility	1.209	1.092	0.776	1.976	1.839	1.180	1.676	1.629	0.839
$T = 120$	Mean	0.567	0.529	0.449	0.867	0.808	0.636	0.887	0.848	0.639
	Vol	0.326	0.308	0.244	0.317	0.281	0.241	0.489	0.479	0.299
	Utility	0.792	0.728	0.583	1.081	0.981	0.768	1.245	1.188	0.793
$T = 240$	Mean	0.425	0.401	0.364	0.638	0.593	0.504	0.650	0.629	0.508
	Vol	0.238	0.230	0.206	0.196	0.168	0.157	0.219	0.212	0.165
	Utility	0.615	0.573	0.504	0.848	0.758	0.642	0.773	0.744	0.586
Three-fund rule										
$T = 60$	Mean	0.797	0.740	0.592	1.452	1.362	0.900	1.076	1.050	0.613
	Vol	0.554	0.515	0.347	0.778	0.735	0.474	0.952	0.926	0.425
	Utility	1.325	1.186	0.809	2.299	2.100	1.218	2.033	1.942	0.817
$T = 120$	Mean	0.531	0.493	0.422	0.859	0.797	0.634	0.863	0.813	0.610
	Vol	0.276	0.267	0.217	0.343	0.291	0.249	0.600	0.575	0.320
	Utility	0.759	0.693	0.554	1.142	1.010	0.787	1.445	1.334	0.797
$T = 240$	Mean	0.395	0.381	0.347	0.562	0.520	0.456	0.487	0.472	0.393
	Vol	0.183	0.191	0.177	0.169	0.146	0.139	0.224	0.213	0.162
	Utility	0.566	0.541	0.479	0.751	0.670	0.581	0.676	0.639	0.491

		25SBTM			**25OPINV**			**30IND**		
		Normal	t	Elliptical	Normal	t	Elliptical	Normal	t	Elliptical
Two-fund rule										
$T = 60$	Mean	1.093	1.063	0.672	1.196	1.165	0.706	0.719	0.704	0.553
	Vol	0.771	0.767	0.426	0.906	0.897	0.510	0.637	0.610	0.459
	Utility	1.662	1.620	0.860	1.928	1.880	0.952	1.044	1.006	0.731
$T = 120$	Mean	0.840	0.801	0.582	1.099	1.054	0.717	0.442	0.432	0.388
	Vol	0.542	0.506	0.338	0.567	0.545	0.344	0.332	0.324	0.285
	Utility	1.269	1.176	0.769	1.562	1.477	0.908	0.566	0.550	0.481
$T = 240$	Mean	0.679	0.645	0.523	0.809	0.780	0.615	0.321	0.316	0.293
	Vol	0.484	0.451	0.336	0.630	0.618	0.427	0.203	0.199	0.183
	Utility	1.125	1.036	0.755	1.474	1.409	0.956	0.374	0.367	0.337
Three-fund rule										
$T = 60$	Mean	1.075	1.021	0.635	1.108	1.074	0.644	0.781	0.750	0.563
	Vol	0.898	0.881	0.450	1.036	1.012	0.531	0.769	0.726	0.513
	Utility	1.923	1.820	0.851	2.165	2.068	0.921	1.304	1.219	0.797
$T = 120$	Mean	0.757	0.719	0.538	0.851	0.809	0.593	0.349	0.338	0.309
	Vol	0.581	0.541	0.344	0.565	0.524	0.305	0.457	0.437	0.358
	Utility	1.298	1.181	0.739	1.397	1.270	0.765	0.629	0.588	0.470
$T = 240$	Mean	0.610	0.586	0.486	0.631	0.606	0.503	0.258	0.253	0.235
	Vol	0.461	0.434	0.330	0.549	0.537	0.385	0.226	0.223	0.199
	Utility	1.065	0.983	0.719	1.272	1.196	0.823	0.383	0.366	0.319

Table C.3: Out-of-Sample Performance of the Combination of the Sample GMV Portfolio with the Risk-Free Asset.

This table reports the annualized gross and net-of-cost out-of-sample utility (EU), and the mean value of the combination coefficient, for the scaled GMV portfolio in Section VI.C.2 across the six datasets described in Section VI.B. The largest EU in each case is depicted in bold. See the notes of Table 3 in the main body of the paper for details.

		Normal (exact)	t (asymp)	t (exact)	Elliptical (asymp)	Elliptical (exact)	Cross validation
A) 10MOM dataset							
$T = 60$	Gross EU	0.041	0.043	0.068	0.108	**0.115**	0.090
	Net EU	0.000	0.003	0.031	0.079	**0.088**	0.054
	Mean $\hat{c}$	0.648	0.609	0.560	0.419	0.380	0.236
$T = 120$	Gross EU	0.124	0.131	0.136	0.145	**0.147**	0.121
	Net EU	0.105	0.113	0.119	0.130	**0.133**	0.102
	Mean $\hat{c}$	0.816	0.755	0.724	0.599	0.570	0.374
$T = 240$	Gross EU	0.078	0.069	0.072	0.077	**0.079**	0.055
	Net EU	0.068	0.059	0.062	0.069	**0.071**	0.045
	Mean $\hat{c}$	0.906	0.844	0.825	0.739	0.720	0.524
B) 16ANOM dataset							
$T = 60$	Gross EU	−0.270	−0.303	−0.209	0.028	**0.061**	0.029
	Net EU	−0.350	−0.377	−0.282	−0.030	**0.007**	−0.033
	Mean $\hat{c}$	0.494	0.480	0.438	0.288	0.260	0.090
$T = 120$	Gross EU	0.070	0.110	0.120	0.123	**0.129**	0.090
	Net EU	0.037	0.079	0.090	0.100	**0.106**	0.066
	Mean $\hat{c}$	0.727	0.677	0.650	0.491	0.468	0.200
$T = 240$	Gross EU	0.262	0.281	**0.283**	0.268	0.267	0.207
	Net EU	0.242	0.263	**0.265**	0.253	0.252	0.190
	Mean $\hat{c}$	0.859	0.811	0.795	0.660	0.644	0.470
C) 25SBETA dataset							
$T = 60$	Gross EU	−0.148	−0.242	−0.113	0.154	**0.155**	0.053
	Net EU	−0.293	−0.387	−0.251	0.075	**0.084**	−0.045
	Mean $\hat{c}$	0.303	0.315	0.280	0.140	0.123	0.077
$T = 120$	Gross EU	−0.178	−0.156	−0.124	0.054	**0.059**	−0.070
	Net EU	−0.231	−0.206	−0.174	0.018	**0.023**	−0.104
	Mean $\hat{c}$	0.604	0.566	0.543	0.340	0.324	0.128
$T = 240$	Gross EU	0.141	0.151	0.153	**0.169**	0.168	0.055
	Net EU	0.111	0.122	0.125	**0.148**	0.147	0.031
	Mean $\hat{c}$	0.791	0.747	0.733	0.543	0.530	0.275

Table C.3: Out-of-Sample Performance of the Combination of the Sample GMV Portfolio with the Risk-Free Asset (continued).

		Normal (exact)	t (asymp)	t (exact)	Elliptical (asymp)	Elliptical (exact)	Cross validation
D) 25SBTM dataset							
$T = 60$	Gross EU	−0.021	−0.072	0.020	**0.181**	0.180	0.128
	Net EU	−0.150	−0.200	−0.103	0.107	**0.114**	0.030
	Mean $\hat{c}$	0.303	0.312	0.277	0.140	0.123	0.077
$T = 120$	Gross EU	0.221	0.240	0.251	**0.268**	0.266	0.238
	Net EU	0.157	0.179	0.192	**0.227**	0.226	0.212
	Mean $\hat{c}$	0.604	0.559	0.536	0.335	0.319	0.510
$T = 240$	Gross EU	0.204	0.215	0.219	**0.237**	0.215	0.227
	Net EU	0.168	0.182	0.186	**0.211**	0.166	0.198
	Mean $\hat{c}$	0.791	0.721	0.707	0.522	0.192	0.382
E) 25OPINV dataset							
$T = 60$	Gross EU	−0.090	−0.144	−0.042	0.147	**0.151**	0.045
	Net EU	−0.242	−0.297	−0.187	0.063	**0.075**	−0.063
	Mean $\hat{c}$	0.303	0.319	0.284	0.138	0.122	0.082
$T = 120$	Gross EU	0.160	0.203	0.217	**0.246**	0.242	0.161
	Net EU	0.088	0.133	0.149	**0.201**	0.199	0.099
	Mean $\hat{c}$	0.604	0.572	0.549	0.336	0.320	0.223
$T = 240$	Gross EU	0.245	0.266	**0.269**	**0.269**	**0.269**	0.243
	Net EU	0.203	0.225	0.229	**0.239**	**0.239**	0.206
	Mean $\hat{c}$	0.791	0.746	0.732	0.537	0.525	0.449
F) 30IND dataset							
$T = 60$	Gross EU	−0.119	−0.140	−0.080	0.009	0.022	**0.034**
	Net EU	−0.194	−0.221	−0.152	−0.043	−0.023	**−0.013**
	Mean $\hat{c}$	0.217	0.233	0.201	0.138	0.119	0.047
$T = 120$	Gross EU	0.058	0.069	0.079	0.104	**0.107**	0.079
	Net EU	0.029	0.041	0.052	0.083	**0.087**	0.056
	Mean $\hat{c}$	0.541	0.504	0.482	0.360	0.344	0.170
$T = 240$	Gross EU	0.041	0.044	0.047	0.061	**0.063**	0.025
	Net EU	0.025	0.030	0.033	0.049	**0.051**	0.011
	Mean $\hat{c}$	0.754	0.698	0.685	0.561	0.549	0.374

Table C.4: Out-of-Sample Performance with Daily Data.

This table reports the gross and net-of-cost annualized out-of-sample utility (EU), and the mean value of the combination coefficients, for the scaled GMV portfolio, the two-fund rule, and the three-fund rules across the six datasets described in Section VI.B. We use sample sizes of $T = 5$, 10, and 20 years of daily returns. The largest EU in each case is depicted in bold. See the notes of Table 3 in the main body of the paper for details.

		Normal (exact)	t (asymp)	Elliptical (asymp)	Normal (exact)	t (asymp)	Elliptical (asymp)
		A) 10MOM dataset ($\hat{\nu} = 3.42$)			**B) 25SBTM dataset** ($\hat{\nu} = 2.99$)		
	Scaled GMV portfolio						
$T = 5$ years	Gross EU	2.348	2.400	**2.582**	3.085	3.220	**3.894**
	Net EU	2.232	2.285	**2.472**	2.835	2.973	**3.657**
	Mean $\hat{c}$	0.982	0.953	0.896	0.958	0.936	0.810
$T = 10$ years	Gross EU	2.267	2.312	**2.381**	3.665	3.701	**4.154**
	Net EU	2.190	2.236	**2.306**	3.489	3.526	**3.983**
	Mean $\hat{c}$	0.991	0.966	0.932	0.979	0.958	0.871
$T = 20$ years	Gross EU	2.328	2.367	**2.446**	3.045	3.077	**3.441**
	Net EU	2.258	2.297	**2.377**	2.849	2.882	**3.254**
	Mean $\hat{c}$	0.995	0.971	0.959	0.990	0.963	0.918
	Two-fund rule						
$T = 5$ years	Gross EU	2.159	2.200	**2.401**	2.843	2.905	**3.687**
	Net EU	1.968	2.011	**2.220**	2.500	2.563	**3.352**
	Mean $\hat{c}$	0.541	0.536	0.508	0.459	0.456	0.407
$T = 10$ years	Gross EU	2.312	2.350	**2.438**	3.424	3.454	**3.970**
	Net EU	2.192	2.231	**2.321**	3.157	3.188	**3.710**
	Mean $\hat{c}$	0.586	0.579	0.559	0.518	0.515	0.473
$T = 20$ years	Gross EU	2.464	2.508	**2.606**	2.568	2.600	**3.024**
	Net EU	2.359	2.403	**2.503**	2.299	2.332	**2.765**
	Mean $\hat{c}$	0.681	0.675	0.660	0.639	0.637	0.603
	Three-fund rule						
$T = 5$ years	Gross EU	2.209	2.267	**2.479**	2.814	2.956	**3.769**
	Net EU	2.039	2.098	**2.317**	2.490	2.635	**3.453**
	Mean $\hat{c}_1$	0.342	0.340	0.326	0.281	0.280	0.259
	Mean $\hat{c}_2/\hat{\mu}_g$	0.640	0.613	0.570	0.678	0.657	0.551
$T = 10$ years	Gross EU	2.350	2.400	**2.477**	3.684	3.725	**4.236**
	Net EU	2.241	2.292	**2.371**	3.435	3.478	**3.993**
	Mean $\hat{c}_1$	0.409	0.407	0.395	0.358	0.356	0.335
	Mean $\hat{c}_2/\hat{\mu}_g$	0.581	0.558	0.537	0.621	0.602	0.536
$T = 20$ years	Gross EU	2.514	2.560	**2.651**	2.875	2.911	**3.326**
	Net EU	2.417	2.464	**2.557**	2.615	2.653	**3.077**
	Mean $\hat{c}_1$	0.527	0.524	0.514	0.488	0.486	0.466
	Mean $\hat{c}_2/\hat{\mu}_g$	0.469	0.446	0.445	0.502	0.477	0.451

Table C.4: Out-of-Sample Performance with Daily Data (continued).

		Normal (exact)	t (asymp)	Elliptical (asymp)	Normal (exact)	t (asymp)	Elliptical (asymp)
		C) 25OPINV dataset ($\hat{v} = 6.27$)			**D) 30IND dataset** ($\hat{v} = 3.86$)		
	Scaled GMV portfolio						
$T = 5$ years	Gross EU	5.951	6.064	**7.248**	2.499	2.685	**3.453**
	Net EU	5.511	5.627	**6.829**	2.291	2.479	**3.256**
	Mean $\hat{c}$	0.958	0.944	0.810	0.951	0.923	0.822
$T = 10$ years	Gross EU	4.928	4.994	**5.871**	2.771	2.906	**3.243**
	Net EU	4.593	4.660	**5.550**	2.643	2.778	**3.119**
	Mean $\hat{c}$	0.979	0.968	0.866	0.975	0.952	0.883
$T = 20$ years	Gross EU	−0.539	−0.470	**0.430**	3.057	3.167	**3.461**
	Net EU	−0.855	−0.785	**0.132**	2.943	3.054	**3.350**
	Mean $\hat{c}$	0.990	0.982	0.912	0.988	0.967	0.927
	Two-fund rule						
$T = 5$ years	Gross EU	5.758	5.835	**7.012**	2.100	2.246	**3.007**
	Net EU	5.157	5.236	**6.437**	1.790	1.936	**2.708**
	Mean $\hat{c}$	0.621	0.617	0.540	0.477	0.472	0.435
$T = 10$ years	Gross EU	4.725	4.766	**5.671**	2.650	2.739	**3.059**
	Net EU	4.300	4.343	**5.260**	2.452	2.542	**2.867**
	Mean $\hat{c}$	0.681	0.677	0.612	0.530	0.525	0.496
$T = 20$ years	Gross EU	−0.692	−0.633	**0.300**	2.699	2.789	**3.084**
	Net EU	-1.068	-1.008	**−0.059**	2.548	2.638	**2.936**
	Mean $\hat{c}$	0.750	0.747	0.695	0.600	0.595	0.574
	Three-fund rule						
$T = 5$ years	Gross EU	5.866	5.982	**7.296**	2.217	2.412	**3.247**
	Net EU	5.339	5.458	**6.786**	1.949	2.145	**2.989**
	Mean $\hat{c}_1$	0.313	0.312	0.283	0.228	0.228	0.218
	Mean $\hat{c}_2/\hat{\mu}_g$	0.645	0.632	0.527	0.722	0.695	0.603
$T = 10$ years	Gross EU	4.718	4.789	**5.752**	2.754	2.888	**3.226**
	Net EU	4.331	4.403	**5.377**	2.594	2.729	**3.070**
	Mean $\hat{c}_1$	0.361	0.360	0.329	0.200	0.200	0.195
	Mean $\hat{c}_2/\hat{\mu}_g$	0.618	0.608	0.537	0.775	0.752	0.689
$T = 20$ years	Gross EU	−0.954	−0.882	**0.084**	2.938	3.049	**3.346**
	Net EU	-1.308	-1.235	**−0.253**	2.812	2.923	**3.223**
	Mean $\hat{c}_1$	0.477	0.476	0.452	0.159	0.158	0.157
	Mean $\hat{c}_2/\hat{\mu}_g$	0.513	0.507	0.461	0.829	0.809	0.770

D. Proofs of Theoretical Results

Proof of Proposition 1

The expected out-of-sample utility of the two-fund rule $\hat{w}_{2f}(c)$ is

$$\mathbb{E}[U(\hat{w}_{2f}(c))] = \frac{c}{\gamma}\tilde{\mu}_1 - \frac{c^2}{2\gamma}\tilde{\sigma}_1^2, \tag{A3}$$

which yields the optimal $c^\star$ in (21). The expected out-of-sample utility of the three-fund rule $\hat{w}_{3f}(c_1, c_2)$ is

$$\mathbb{E}[U(\hat{w}_{3f}(c_1,c_2))] = \frac{c_1}{\gamma}\tilde{\mu}_1 + \frac{c_2}{\gamma}\tilde{\mu}_2 - \frac{c_1^2}{2\gamma}\tilde{\sigma}_1^2 - \frac{c_2^2}{2\gamma}\tilde{\sigma}_2^2 - \frac{c_1c_2}{\gamma}\tilde{\sigma}_{12}, \tag{A4}$$

which yields the optimal $(c_1^\star, c_2^\star)$ in (22) and completes the proof.

Proof of Proposition 2

Part 1. Because $\hat{\mu}$ and $\hat{\Sigma}$ are asymptotically unbiased for a fixed N, the sample mean-variance portfolio $\hat{w}$ is asymptotically unbiased too. To find its asymptotic covariance matrix, we write $\hat{w}$ as a function of $\hat{\mu}$ and $\text{vec}(\hat{\Sigma}^{-1})$ as

$$\hat{w} = \frac{1}{\gamma}(\hat{\mu}^\top \otimes I_N)\text{vec}(\hat{\Sigma}^{-1}). \tag{A5}$$

Moreover, the derivative of $\text{vec}(\hat{\Sigma}^{-1})$ with respect to $\text{vec}(\hat{\Sigma})$ is

$$\text{(A6)} \qquad \frac{\partial \text{vec}(\hat{\Sigma}^{-1})}{\partial \text{vec}(\hat{\Sigma})^\top} = -\hat{\Sigma}^{-1} \otimes \hat{\Sigma}^{-1}.$$

Therefore, using the delta method, we can find the asymptotic covariance matrix of $\hat{w}$ from the asymptotic covariance matrix of $(\hat{\mu}, \text{vec}(\hat{\Sigma}))$, which from Muirhead (1982, p.82, 89) is

$$\text{(A7)} \qquad \text{Avar}\begin{bmatrix} \hat{\mu} \\ \text{vec}(\hat{\Sigma}) \end{bmatrix} = \begin{bmatrix} \Sigma & 0_{N\times N^2} \\ 0_{N^2\times N} & (1+\kappa)(I_{N^2}+K_N)(\Sigma\otimes\Sigma)+\kappa\,\text{vec}(\Sigma)\text{vec}(\Sigma)^\top \end{bmatrix},$$

where K_N is an $N^2 \times N^2$ commutation matrix such that $K_N \text{vec}(A) = \text{vec}(A^\top)$ for an $N \times N$ matrix A. Specifically, given (A5)–(A6), we have

$$\text{(A8)} \qquad \frac{\partial \hat{w}}{\partial \hat{\mu}^\top} = \frac{1}{\gamma}\hat{\Sigma}^{-1},$$

$$\text{(A9)} \qquad \frac{\partial \hat{w}}{\partial \text{vec}(\hat{\Sigma})^\top} = -\frac{1}{\gamma}(\hat{\Sigma}^{-1}\hat{\mu}) \otimes \hat{\Sigma}^{-1},$$

and therefore the asymptotic covariance matrix of $\hat{w}$ is

$$\text{(A10)} \qquad \text{Avar}[\hat{w}] = \frac{1}{\gamma^2}\left[\Sigma^{-1}, -(\Sigma^{-1}\mu)\otimes\Sigma^{-1}\right]\text{Avar}\begin{bmatrix} \hat{\mu} \\ \text{vec}(\hat{\Sigma}) \end{bmatrix}\left[\Sigma^{-1}, -(\Sigma^{-1}\mu)\otimes\Sigma^{-1}\right]^\top,$$

which after simplification corresponds to the desired result in (23).

Part 2. Given that $\hat{w}$ is asymptotically unbiased as shown in part 1, $\tilde{\mu}_p$ and $\tilde{\sigma}_p^2$ are asymptotically unbiased too. To find the asymptotic covariance matrix of $\tilde{\mu}_p$ and $\tilde{\sigma}_p^2$, we use the delta method. Let $h(w) = [w^\top \mu, w^\top \Sigma w]^\top$. Then, the Jacobian of h evaluated at $w^\star$ is

$\nabla h(w^\star) = [\mu^\top, 2w^{\star\top}\Sigma]^\top = [\mu^\top, 2\mu^\top/\gamma]^\top$. Therefore, the asymptotic covariance matrix of $\tilde{\mu}_p$ and $\tilde{\sigma}_p^2$ is $\text{Avar}[\tilde{\mu}_p, \tilde{\sigma}_p^2] = \nabla h(w^\star)^\top \text{Avar}[\hat{w}] \nabla h(w^\star)$, where $\text{Avar}[\hat{w}]$ is given by (**??**), which corresponds to the desired result in (24).

Part 3. Given $\hat{w} = \frac{1}{\gamma}\hat{\Sigma}^{-1}\hat{\mu}$ and $U(\hat{w}) = \hat{w}^\top \mu - \hat{w}^\top \Sigma \hat{w}$, we have

$$D = \left.\frac{\partial U(\hat{w})}{\partial \hat{w}}\right|_{\hat{w}=w^*} = \mu - \gamma\Sigma w^\star = 0_N, \tag{A11}$$

$$H = \left.\frac{\partial^2 U(\hat{w})}{\partial \hat{w}\partial \hat{w}^\top}\right|_{\hat{w}=w^*} = -\gamma\Sigma. \tag{A12}$$

Because $D = 0_N$, it holds that

$$T[U(\hat{w}) - U(w^\star)] \xrightarrow{d} \sum_{i=1}^{N} \lambda_i X_i, \tag{A13}$$

where the X_i's are independent χ_1^2 random variables and the λ_i's are the eigenvalues of $HS/2$, where S is the asymptotic covariance matrix of $\hat{w}$ in (23):

$$S = \frac{1}{\gamma^2}\Big((1 + (1+\kappa)\theta^2)\Sigma^{-1} + (1+2\kappa)\Sigma^{-1}\mu\mu^\top\Sigma^{-1}\Big). \tag{A14}$$

The matrix $HS/2$ is

$$\frac{1}{2}HS = -\frac{1+(1+\kappa)\theta^2}{2\gamma}\left(I_N + \frac{1+2\kappa}{1+(1+\kappa)\theta^2}\mu\mu^\top\Sigma^{-1}\right). \tag{A15}$$

All the eigenvalues of I_N are one, and $\mu\mu^\top\Sigma^{-1}$ has $N-1$ zero eigenvalues and one eigenvalue equal to the trace of $\mu\mu^\top\Sigma^{-1}$, i.e., θ^2. Therefore, we have $\lambda_1 = \cdots = \lambda_{N-1} = -\frac{1+(1+\kappa)\theta^2}{2\gamma}$ and

$\lambda_N = -\frac{1+(1+\kappa)\theta^2}{2\gamma}\left(1+\frac{(1+2\kappa)\theta^2}{1+(1+\kappa)\theta^2}\right) = -\frac{1+(2+3\kappa)\theta^2}{2\gamma}$, which yields the desired result in (25).

Part 4. Let the risk-aversion coefficient $\gamma = 1$ for notational simplicity, which is without loss of generality because it is clear that $\tilde{\mu}_p$, $\tilde{\sigma}_p^2$, and $U(\hat{w})$ are proportional to $1/\gamma$, $1/\gamma^2$, and $1/(2\gamma)$, respectively. Let $\phi = [\mu^\top, w^{\star\top}]^\top$ and $\hat{\phi} = [\hat{\mu}^\top, \hat{w}^\top]^\top$. Note that $\hat{\phi}$ can be written as the generalized-method-of-moments estimator of ϕ based on the following moment conditions:

$$\mathbb{E}[g_t(\phi)] = \mathbb{E}\begin{bmatrix} r_t - \mu \\ C_t w^\star - \mu \end{bmatrix} = 0_{2N}, \tag{A16}$$

where $C_t = (r_t - \mu)(r_t - \mu)^\top$.

We derive the first-order bias of $\hat{\phi}$ by using a stochastic expansion of $\hat{\phi}$ based on the results of Bao and Ullah (2007, 2009), which suggest

$$\hat{\phi} = \phi + a_{-1/2} + a_{-1} + a_{-3/2} + O_p(T^{-2}), \tag{A17}$$

where

$$a_{-1/2} = -\mathbb{E}[H_1]^{-1}\bar{g}, \tag{A18}$$

$$a_{-1} = -\mathbb{E}[H_1]^{-1}V_1 a_{-1/2} - \frac{1}{2}\mathbb{E}[H_1]^{-1}\mathbb{E}[H_2](a_{-1/2}\otimes a_{-1/2}), \tag{A19}$$

$$\begin{aligned} a_{-3/2} = &-\mathbb{E}[H_1]^{-1}V_1 a_{-1} - \frac{1}{2}\mathbb{E}[H_1]^{-1}V_2(a_{-1/2}\otimes a_{-1/2}) \\ &-\frac{1}{2}\mathbb{E}[H_1]^{-1}\mathbb{E}[H_2](a_{-1/2}\otimes a_{-1} + a_{-1}\otimes a_{-1/2}) \\ &-\frac{1}{6}\mathbb{E}[H_1]^{-1}\mathbb{E}[H_3](a_{-1/2}\otimes a_{-1/2}\otimes a_{-1/2}), \end{aligned} \tag{A20}$$

with $\bar{g} = \frac{1}{T}\sum_{t=1}^{T} g_t(\phi)$, $H_i = \nabla^i \bar{g}$, and $V_i = H_i - \mathbb{E}[H_i]$.[1]

We now provide explicit expressions of $a_{-1/2}$ and a_{-1}. For $a_{-3/2}$, we can show that its expectation is $O(T^{-2})$.[2] For $a_{-1/2}$, we have

$$H_1 = \begin{bmatrix} -I_N & 0_{N\times N} \\ -z_t I_N - (r_t-\mu)w^{\star\top} - I_N & C_t \end{bmatrix}, \tag{A21}$$

where $z_t = (r_t-\mu)^\top w^\star$, and thus,

$$\mathbb{E}[H_1] = \begin{bmatrix} -I_N & 0_{N\times N} \\ -I_N & \Sigma \end{bmatrix}, \tag{A22}$$

$$\mathbb{E}[H_1]^{-1} = \begin{bmatrix} -I_N & 0_{N\times N} \\ -\Sigma^{-1} & \Sigma^{-1} \end{bmatrix}. \tag{A23}$$

It follows that $a_{-1/2}$ is equal to

$$a_{-1/2} = \frac{1}{T}\sum_{t=1}^{T} \begin{bmatrix} r_t - \mu \\ a_t \end{bmatrix}, \tag{A24}$$

where

$$a_t = \Sigma^{-1}(r_t-\mu)(1-z_t) + w^*, \tag{A25}$$

[1] $\nabla^i \bar{g}$ is the matrix of i-th order partial derivative of $\bar{g}(\phi)$ and is obtained recursively. Specifically, If $\bar{g}(\phi)$ is a k-vector function of ϕ, the j-th element of the l-th row of $A_i \equiv \nabla^i \bar{g}$ (a $k \times k^i$ matrix) is the $1\times k$ vector $a^i_{lj} = \partial a^{i-1}_{lj}/\partial \phi^\top$.

[2] An explicit expression of $\mathbb{E}[a_{-3/2}]$ is available upon request.

and it is obvious that $\mathbb{E}[a_{-1/2}] = 0_{2N}$.

For a_{-1}, we have from (A21) and (A22) that

$$V_1 = \frac{1}{T}\sum_{t=1}^{T}\begin{bmatrix} 0_{N\times N} & 0_{N\times N} \\ -z_t I_N - (r_t-\mu)w^{\star\top} & C_t - \Sigma \end{bmatrix}. \tag{A26}$$

Moreover,

$$H_2 = \frac{1}{T}\sum_{t=1}^{T}\begin{bmatrix} 0_{N\times 2N^2} & 0_{N\times 2N^2} \\ A_t & B_t \end{bmatrix}, \tag{A27}$$

where

$$A_t = I_N \otimes [w^{\star\top},\ -(r_t-\mu)^\top] + w^{\star\top}\otimes D_1^\top - (r_t-\mu)\mathrm{vec}(D_2)^\top, \tag{A28}$$

$$B_t = -(r_t-\mu)^\top \otimes D_1^\top - (r_t-\mu)\mathrm{vec}(D_1)^\top, \tag{A29}$$

with

$$D_1 = \begin{bmatrix} I_N \\ 0_{N\times N} \end{bmatrix}, \tag{A30}$$

$$D_2 = \begin{bmatrix} 0_{N\times N} \\ I_N \end{bmatrix}. \tag{A31}$$

Therefore,

$$\text{(A32)} \qquad \mathbb{E}[H_2] = \begin{bmatrix} 0_{N\times 2N^2} & 0_{N\times 2N^2} \\ I_N \otimes [w^{\star\top},\, 0_N^\top] + w^{\star\top} \otimes D_1^\top & 0_{N\times 2N^2} \end{bmatrix}.$$

It follows that a_{-1} is equal to

$$\text{(A33)} \qquad a_{-1} = \frac{1}{T^2} \sum_{s=1}^{T} \sum_{t=1}^{T} \begin{bmatrix} 0_N \\ z_s \Sigma^{-1}(r_t - \mu) - \Sigma^{-1}(C_s - \Sigma) a_t \end{bmatrix}.$$

Using the fact that r_s is independent of r_t if $s \neq t$, $\mathbb{E}[z_t] = 0$, $\mathbb{E}[a_t] = 0_N$, and $\mathbb{E}[C_t - \Sigma] = 0_{N\times N}$, the expectation of a_{-1} is

$$\text{(A34)} \qquad \mathbb{E}[a_{-1}] = \frac{1}{T} \begin{bmatrix} 0_N \\ \mathbb{E}[z_t \Sigma^{-1}(r_t - \mu)] - \mathbb{E}[\Sigma^{-1}(C_t - \Sigma) a_t] \end{bmatrix}.$$

Let $y_t = \Sigma^{-1/2}(r_t - \mu)$. Then,

$$\text{(A35)} \qquad \mathbb{E}[z_t \Sigma^{-1}(r_t - \mu)] = \mathbb{E}[y_t y_t^\top] w^\star = w^\star,$$

$$\text{(A36)} \qquad \mathbb{E}[\Sigma^{-1}(C_t - \Sigma) a_t] = \mathbb{E}[y_t y_t^\top] w^\star - \mathbb{E}[(y_t^\top y_t) y_t y_t^\top] w^\star = [1 - (N+2)(1+\kappa)] w^\star,$$

where we use the fact that $\mathbb{E}[(y_t^\top y_t) y_t y_t^\top] = (N+2)(1+\kappa) I_N$ under the multivariate elliptical distribution assumption. Therefore,

$$\text{(A37)} \qquad \mathbb{E}[a_{-1}] = \begin{bmatrix} 0_N \\ \frac{1}{T}(N+2)(1+\kappa) w^\star \end{bmatrix}.$$

Using $\mathbb{E}[a_{-1/2}] = 0_{2N}$ and $\mathbb{E}[a_{-1}]$ in (A37), the first-order bias of $\hat{w}$ is

$$\mathbb{E}[\hat{w}] - w^\star = \frac{(N+2)(1+\kappa)w^\star}{T} + O(T^{-2}). \tag{A38}$$

It follows that

$$\mathbb{E}[\tilde{\mu}_p] - \mu_p = \frac{(N+2)(1+\kappa)\mu_p}{T} + O(T^{-2}), \tag{A39}$$

which corresponds to the desired result in (26) after adding back $1/\gamma$.

Turning to the first-order bias of $\tilde{\sigma}_p^2$, we use (A17) to obtain

$$\mathbb{E}[\tilde{\sigma}_p^2] - \sigma_p^2 = 2w^{\star\top}\Sigma\mathbb{E}[\tilde{a}_{-1}] + \mathbb{E}[\tilde{a}_{-1/2}^\top \Sigma \tilde{a}_{-1/2}] + O(T^{-2}), \tag{A40}$$

where $\tilde{a}_{-1/2}$ and $\tilde{a}_{-1}$ are the last N elements of $a_{-1/2}$ and a_{-1}, respectively. Given $\mathbb{E}[\tilde{a}_{-1}]$ in (A37), we have

$$2w^{\star\top}\Sigma\mathbb{E}[\tilde{a}_{-1}] = \frac{2(N+2)(1+\kappa)\theta^2}{T}. \tag{A41}$$

Moreover, given $\tilde{a}_{-1/2}$ in (A24), we have

$$\tilde{a}_{-1/2}^\top \Sigma \tilde{a}_{-1/2} = \frac{1}{T^2}\sum_{s=1}^{T}\sum_{t=1}^{T} a_s^\top \Sigma a_t. \tag{A42}$$

Therefore,

$$\mathbb{E}[\tilde{a}_{-1/2}^\top \Sigma \tilde{a}_{-1/2}] = \frac{1}{T}\mathbb{E}[a_t^\top \Sigma a_t]$$

$$= \frac{1}{T}\Big(\theta^2 + 2\mathbb{E}[(1-z_t)(r_t-\mu)^\top \Sigma^{-1}\mu] + \mathbb{E}[(1-z_t)^2(r_t-\mu)^\top \Sigma^{-1}(r_t-\mu)]\Big). \tag{A43}$$

It holds that

$$\mathbb{E}[(1-z_t)(r_t-\mu)^\top \Sigma^{-1}\mu] = -\theta^2, \tag{A44}$$

$$\mathbb{E}[(1-z_t)^2(r_t-\mu)^\top \Sigma^{-1}(r_t-\mu)] = \mathbb{E}[y_t^\top y_t] + \mathbb{E}[\mu^\top \Sigma^{-1/2}(y_t^\top y_t) y_t y_t^\top \Sigma^{-1/2}\mu]$$

$$= N + (N+2)(1+\kappa)\theta^2, \tag{A45}$$

and thus,

$$\mathbb{E}[\tilde{a}_{-1/2}^\top \Sigma \tilde{a}_{-1/2}] = \frac{N + [(N+2)(1+\kappa)-1]\theta^2}{T}. \tag{A46}$$

It follows that

$$\mathbb{E}[\tilde{\sigma}_p^2] - \sigma_p^2 = \frac{N + [3(N+2)(1+\kappa)-1]\theta^2}{T} + O(T^{-2}), \tag{A47}$$

which corresponds to the desired result in (27) after adding back $1/\gamma^2$. Finally, the first-order bias of $U(\hat{w})$ in (28) is directly obtained from (26)–(27). This completes the proof.

Proof of Proposition 3

Part 1. This result is a direct consequence of El Karoui (2010, equation (9)) after defining the parameter η as $(1-\rho)\mathfrak{s}$, where $\mathfrak{s}$ is defined in El Karoui (2010, equation (4)). El Karoui (2010) shows in p. 3506 that $\mathfrak{s} \geq 1/(1-\rho)$, and thus $\eta \geq 1$. Finally, we show in part 4 of this proposition that $\lim_{\rho\to 1}\eta = \mathbb{E}[1/\tau_t]$.

Part 2. From El Karoui (2013, equation (3.4)), we have

$$\tilde{\sigma}_p^2 = \frac{1}{\gamma^2}\hat{\mu}^\top \hat{\Sigma}^{-1}\Sigma\hat{\Sigma}^{-1}\hat{\mu} \xrightarrow{p} \frac{1}{\gamma^2}\mu^\top \hat{\Sigma}^{-1}\Sigma\hat{\Sigma}^{-1}\mu + \frac{\eta\rho}{\gamma^2(1-\rho)^3}. \tag{A48}$$

Moreover, using the last equation in p. 748 of El Karoui (2013), we obtain

$$\mu^\top \hat{\Sigma}^{-1}\Sigma\hat{\Sigma}^{-1}\mu \xrightarrow{p} \frac{\varphi\theta^2}{(1-\rho)^3}, \tag{A49}$$

where φ is defined as $(1-\rho)^3\xi$ with ξ defined in El Karoui (2013, equation (3.2)). El Karoui (2013, fact 3.1) shows that $\xi \geq \mathfrak{s}^2/(1-\rho)$, which in our notation is equivalent to $\varphi \geq \eta^2$. Finally, we show in part 4 of this proposition that $\lim_{\rho\to 1}\varphi = (\mathbb{E}[1/\tau_t])^2$.

Part 3. Equation (33) is a direct consequence of (29)–(31). As shown in part 4 of this proposition, the case of multivariate normally distributed returns corresponds to $\eta = \varphi = 1$. Therefore, the limit of $U(\hat{w})$ in (33) is smaller than that under normality if

$$\frac{2\eta}{1-\rho} - \frac{\varphi}{(1-\rho)^3} \leq \frac{2}{1-\rho} - \frac{1}{(1-\rho)^3}, \tag{A50}$$

which is equivalent to

$$\rho \geq 1-\sqrt{\frac{\varphi-1}{2(\eta-1)}}. \tag{A51}$$

Condition (A51) always holds because the right-hand side of (A51) is negative:

$$1-\sqrt{\frac{\varphi-1}{2(\eta-1)}} \leq 1-\sqrt{\frac{\eta^2-1}{2(\eta-1)}} = 1-\sqrt{\frac{\eta+1}{2}} \leq 0, \tag{A52}$$

where the first and last inequalities hold because $\varphi \geq \eta^2$ and $\eta \geq 1$, respectively.

Part 4. El Karoui (2010, p. 3506) and El Karoui (2013, Fact 3.1) show that $\eta = \varphi = 1$ when returns are multivariate normally distributed. Moreover, when $\rho \to 0$ we recover the fixed N asymptotic regime in Proposition 2, in which case $\eta = \varphi = 1$ too because $\tilde{\mu}_p$ and $\tilde{\sigma}_p^2$ are asymptotically unbiased. Finally, we show that $\lim_{\rho\to 1}\eta = \mathbb{E}[1/\tau_t]$ and $\lim_{\rho\to 1}\varphi = (\mathbb{E}[1/\tau_t])^2$. For η, it is direct from its definition in (30). For φ, we apply L'Hopital's rule to obtain

$$\lim_{\rho\to 1}\varphi = \left(\lim_{\rho\to 1}\frac{2}{\eta^3}\frac{\partial\eta}{\partial\rho} + \mathbb{E}\left[\frac{\tau^2}{(1-\rho+\rho\eta\tau_t)^2}\right] + 2\rho\mathbb{E}\left[\frac{\tau_t^2(1-\tau_t(\eta+\rho\frac{\partial\eta}{\partial\rho}))}{(1-\rho+\rho\eta\tau_t)^3}\right]\right)^{-1}. \tag{A53}$$

We apply implicit differentiation on (30) to obtain

$$\frac{\partial\eta}{\partial\rho} = \frac{\mathbb{E}\left[\frac{1-\eta\tau_t}{(1-\rho+\rho\eta\tau_t)^2}\right]}{\mathbb{E}\left[\frac{\rho\tau_t}{(1-\rho+\rho\eta\tau_t)^2}\right]}, \tag{A54}$$

whose limit as $\rho \to 1$ is

$$\lim_{\rho\to 1}\frac{\partial\eta}{\partial\rho} = \frac{\mathbb{E}[1/\tau_t^2]}{\mathbb{E}[1/\tau_t]} - \mathbb{E}[1/\tau_t]. \tag{A55}$$

Finally, using (A55) and $\lim_{\rho \to 1} \eta = \mathbb{E}[1/\tau_t]$, the limit in (A53) simplifies to $(\mathbb{E}[1/\tau_t])^2$. This completes the proof.

Proof of Proposition 4

Using the results in the proof of Proposition 3, the quantities needed in Proposition 1 to identify the optimal combination coefficients are

$$\tilde{\mu}_1 = \frac{\eta\theta^2}{1-\rho}, \tag{A56}$$

$$\tilde{\mu}_2 = \frac{\eta}{1-\rho}\frac{\theta_g^2}{\mu_g}, \tag{A57}$$

$$\tilde{\sigma}_1^2 = \frac{\varphi\theta^2 + \eta\rho}{(1-\rho)^3}, \tag{A58}$$

$$\tilde{\sigma}_2^2 = \frac{\varphi}{(1-\rho)^3}\frac{\theta_g^2}{\mu_g^2}, \tag{A59}$$

$$\tilde{\sigma}_{12} = \frac{\varphi}{(1-\rho)^3}\frac{\theta_g^2}{\mu_g}. \tag{A60}$$

Plugging (A56)–(A60) into (21)–(22) delivers the optimal two-fund and three-fund combination coefficients in (34)–(36). Finally, these combination coefficients are smaller than those under normally distributed returns because $\varphi \geq \eta^2 \geq \eta$. This completes the proof.

Proof of Proposition 5

Part 1. Given $\tilde{\mu}_1$ and $\tilde{\sigma}_1^2$ in (A56) and (A58), we have from (A3) that the out-of-sample utility of the two-fund rule $\hat{w}_{2f}(c)$ converges to

$$U(\hat{w}_{2f}(c)) \xrightarrow{p} \frac{c}{\gamma}\frac{\eta\theta^2}{1-\rho} - \frac{c^2}{2\gamma}\frac{\varphi\theta^2 + \eta\rho}{(1-\rho)^3}. \tag{A61}$$

Plugging the limit of $c^\star$ in (34) into (A61) yields the desired result in (37). This utility is larger than that under the multivariate normal distribution if and only if

(A62)
$$\frac{\eta\theta^2}{\frac{\varphi}{\eta}\theta^2+\rho} > \frac{\theta^2}{\theta^2+\rho},$$

which is equivalent to the condition $\theta^2 < \rho\eta(\eta-1)/(\varphi-\eta^2)$.

Part 2. Given $\tilde{\mu}_1$, $\tilde{\mu}_2$, $\tilde{\sigma}_1^2$, $\tilde{\sigma}_2^2$, and $\tilde{\sigma}_{12}$ in (A56)–(A60), we have from (A4) that the out-of-sample utility of the three-fund rule $\hat{w}_{3f}(c)$ converges to

(A63)
$$U(\hat{w}_{3f}(c_1,c_2)) \xrightarrow{p} \frac{c_1}{\gamma}\frac{\eta\theta^2}{1-\rho} + \frac{c_2}{\gamma}\frac{\eta}{1-\rho}\frac{\theta_g^2}{\mu_g} - \frac{c_1^2}{2\gamma}\frac{\varphi\theta^2+\eta\rho}{(1-\rho)^3} - \frac{c_2^2}{2\gamma}\frac{\varphi}{(1-\rho)^3}\frac{\theta_g^2}{\mu_g^2} - \frac{c_1c_2}{\gamma}\frac{\varphi}{(1-\rho)^3}\frac{\theta_g^2}{\mu_g}.$$

Plugging the limit of $(c_1^\star, c_2^\star)$ in (35)–(36) into (A63) yields the desired result in (38). This utility is larger than that under the multivariate normal distribution if and only if

(A64)
$$\frac{\eta\psi^2}{\frac{\varphi}{\eta}\psi^2+\rho}\left(1+\frac{\eta\rho\theta_g^2}{\varphi\theta^2\psi^2}\right) > \frac{\psi^2}{\psi^2+\rho}\left(1+\frac{\rho\theta_g^2}{\theta^2\psi^2}\right),$$

which is equivalent to the condition $\eta(\psi^2+\rho)(\theta^2\psi^2+\frac{\eta}{\varphi}\theta_g^2\rho) > (\frac{\varphi}{\eta}\psi^2+\rho)(\theta^2\psi^2+\theta_g^2\rho)$. This completes the proof.

Proof of Proposition 6

Part 1. Given the distribution of τ_t in (2), the expectation in (30) evaluates to

$$\mathbb{E}\left[(1-\rho+\rho\eta\tau_t)^{-1}\right] = \frac{\nu}{2(1-\rho)} e^y E_{\frac{\nu}{2}+1}(y), \tag{A65}$$

where $y = (\nu-2)\rho\eta/[2(1-\rho)]$. Using the recursive relation on the exponential integral,

$$E_{\frac{\nu}{2}+1}(y) = \frac{e^{-y} - yE_{\frac{\nu}{2}}(y)}{\nu/2}. \tag{A66}$$

Therefore, (A65) simplifies to

$$\mathbb{E}\left[(1-\rho+\rho\eta\tau_t)^{-1}\right] = \frac{1-ye^yE_{\frac{\nu}{2}}(y)}{1-\rho}, \tag{A67}$$

and thus condition (30) means that η is the solution to (39). Finally, $\mathbb{E}[1/\tau_t] = \nu/(\nu-2)$, and thus $1 \leq \eta \leq \nu/(\nu-2)$ from Proposition 3.

Part 2. Given the distribution of τ_t in (2), the expectation in (32) evaluates to

$$\mathbb{E}\left[\frac{\tau_t^2}{(1-\rho+\rho\eta\tau_t)^2}\right] = \frac{1}{(1-\rho)^2}\left[-\left(\frac{\nu-2}{2}\right) + \left(\frac{\nu-2}{2}\right)^2\left(1+\frac{\rho\eta}{1-\rho}\right)e^yE_{\frac{\nu}{2}-1}(y)\right], \tag{A68}$$

where $y = (\nu-2)\rho\eta/[2(1-\rho)]$. Using the recursive relation on the exponential integral,

$$E_{\frac{\nu}{2}-1}(y) = \frac{e^{-y} - (\frac{\nu}{2}-1)E_{\frac{\nu}{2}}(y)}{y}. \tag{A69}$$

Moreover, it holds that $e^y E_{\frac{\nu}{2}}(y) = \rho/y$ from (39). Therefore,

$$e^y E_{\frac{\nu}{2}-1}(y) = \frac{1}{y} - \frac{\rho(\frac{\nu}{2}-1)}{y^2}. \tag{A70}$$

Plugging (A70) and $y = \frac{(\nu-2)\rho\eta}{2(1-\rho)}$ into (A68) yields

$$\mathbb{E}\left[\frac{\tau_t^2}{(1-\rho+\rho\eta\tau_t)^2}\right] = \frac{(\nu-2)(\eta-1)}{2\rho\eta^2}. \tag{A71}$$

Plugging (A71) into (32) delivers the desired formula for φ in (40). Finally, $(\mathbb{E}[1/\tau_t])^2 = \nu^2/(\nu-2)^2$, and thus $\eta^2 \leq \varphi \leq \nu^2/(\nu-2)^2$ from Proposition 3. This completes the proof.

Proof of Proposition 7

Given the definition of Y, Λ, M, and the stochastic representation for the multivariate elliptical distribution in (1), we can write the sample mean and covariance matrix as

$$\hat{\mu} = \mu + \frac{1}{T}\Sigma^{1/2}Y^\top\Lambda 1_T, \tag{A72}$$

$$\hat{\Sigma} = \frac{1}{T}\Sigma^{1/2}Y^\top\Lambda M\Lambda Y\Sigma^{1/2}. \tag{A73}$$

Therefore, the out-of-sample mean and variance of the sample mean-variance portfolio $\hat{w}$ are

$$\tilde{\mu}_p = \frac{1}{\gamma}\Big[T\mu^\top\Sigma^{-1/2}(Y^\top\Lambda M\Lambda Y)^{-1}\Sigma^{-1/2}\mu + 1_T^\top\Lambda Y(Y^\top\Lambda M\Lambda Y)^{-1}\Sigma^{-1/2}\mu\Big], \tag{A74}$$

$$\tilde{\sigma}_p^2 = \frac{1}{\gamma^2}\Big[T^2\mu^\top\Sigma^{-1/2}(Y^\top\Lambda M\Lambda Y)^{-2}\Sigma^{-1/2}\mu + 2T1_T^\top\Lambda Y(Y^\top\Lambda M\Lambda Y)^{-2}\Sigma^{-1/2}\mu$$

$$\text{(A75)} \qquad + 1_T^\top \Lambda Y (Y^\top \Lambda M \Lambda Y)^{-2} Y^\top \Lambda 1_T \Big].$$

By symmetry, the expectation of the second terms in (A74)–(A75) is zero because Y has zero mean. Therefore, the expected out-of-sample mean and variance are

$$\text{(A76)} \quad \mathbb{E}[\tilde{\mu}_p] = \frac{1}{\gamma} T \mu^\top \Sigma^{-1/2} \mathbb{E}\Big[(Y^\top \Lambda M \Lambda Y)^{-1}\Big] \Sigma^{-1/2} \mu,$$

$$\text{(A77)} \quad \mathbb{E}[\tilde{\sigma}_p^2] = \frac{1}{\gamma^2} T^2 \mu^\top \Sigma^{-1/2} \mathbb{E}\Big[(Y^\top \Lambda M \Lambda Y)^{-2}\Big] \Sigma^{-1/2} \mu + \frac{1}{\gamma^2} \mathbb{E}\Big[1_T^\top \Lambda Y (Y^\top \Lambda M \Lambda Y)^{-2} Y^\top \Lambda 1_T\Big].$$

By definition of k_3, the second term in (A77) is equal to

$$\text{(A78)} \qquad \frac{1}{\gamma^2} \mathbb{E}\Big[1_T^\top \Lambda Y (Y^\top \Lambda M \Lambda Y)^{-2} Y^\top \Lambda 1_T\Big] = \frac{NT(T-2)}{(T-N-1)(T-N-2)(T-N-4)} \frac{k_3}{\gamma^2}.$$

Moreover, $Y^\top \Lambda M \Lambda Y = T \Sigma^{-1/2} \hat{\Sigma} \Sigma^{-1/2}$, and thus by symmetry $\mathbb{E}[(Y^\top \Lambda M \Lambda Y)^{-1}]$ and $\mathbb{E}[(Y^\top \Lambda M \Lambda Y)^{-2}]$ are both proportional to the identity matrix I_N. If we denote the proportionality constants by a_1 and a_2, then

$$\text{(A79)} \qquad \frac{1}{\gamma} T \mu^\top \Sigma^{-1/2} \mathbb{E}\Big[(Y^\top \Lambda M \Lambda Y)^{-1}\Big] \Sigma^{-1/2} \mu = T a_1 \mu_p,$$

$$\text{(A80)} \qquad \frac{1}{\gamma^2} T^2 \mu^\top \Sigma^{-1/2} \mathbb{E}\Big[(Y^\top \Lambda M \Lambda Y)^{-2}\Big] \Sigma^{-1/2} \mu = T^2 a_2 \sigma_p^2,$$

which proves (44)–(45) because $k_1 = (T-N-2)a_1$ and $k_2 = \frac{(T-N-1)(T-N-2)(T-N-4)}{T-2} a_2$ by definition. Finally, it is known from Kan and Zhou (2007) that when asset returns are multivariate normally distributed, the expectations of $\tilde{\mu}_p$ and $\tilde{\sigma}_p^2$ evaluate to (44) and (45), respectively, with $k_1 = k_2 = k_3 = 1$. This completes the proof.

Proof of Proposition 8

Using the results in the proof of Proposition 7, the quantities needed in Proposition 1 to identify the optimal combination coefficients are

$$\tilde{\mu}_1 = \frac{T}{T-N-2} k_1 \theta^2, \tag{A81}$$

$$\tilde{\mu}_2 = \frac{T}{T-N-2} k_1 \frac{\theta_g^2}{\mu_g}, \tag{A82}$$

$$\tilde{\sigma}_1^2 = \frac{T^2(T-2)}{(T-N-1)(T-N-2)(T-N-4)} \left(k_2 \theta^2 + k_3 \frac{N}{T} \right), \tag{A83}$$

$$\tilde{\sigma}_2^2 = \frac{T^2(T-2)}{(T-N-1)(T-N-2)(T-N-4)} k_2 \frac{\theta_g^2}{\mu_g^2}, \tag{A84}$$

$$\tilde{\sigma}_{12} = \frac{T^2(T-2)}{(T-N-1)(T-N-2)(T-N-4)} k_2 \frac{\theta_g^2}{\mu_g}. \tag{A85}$$

Plugging (A81)–(A85) into (21)–(22) delivers the optimal two-fund and three-fund combination coefficients in (50)–(52). This completes the proof.

Proof of Proposition 9

The proof of this proposition is similar to that for the optimal three-fund combination coefficients in Propositions 4 and 8 when we constrain the combination coefficient on the sample mean-variance portfolio to be $c_1 = 0$.

References

Bao, Y., and A. Ullah. "The second-order bias and mean squared error of estimators in time-series models." *Journal of Econometrics*, 140 (2007), 650–669.

Bao, Y., and A. Ullah. "Higher-order bias and MSE of nonlinear estimators." *Pakistan Journal of Statistics*, 25 (2009), 287–294.

El Karoui, N. "High-dimensionality effects in the Markowitz problem and other quadratic programs with linear constraints: Risk underestimation." *Annals of Statistics*, 38 (2010), 3487–3566.

El Karoui, N. "On the realized risk of high-dimensional Markowitz portfolios." *SIAM Journal on Financial Mathematics*, 4 (2013), 737–783.

Kan, R., and G. Zhou. "Optimal portfolio choice with parameter uncertainty." *Journal of Financial and Quantitative Analysis*, 42 (2007), 621–656.

Muirhead, R. *Aspects of Multivariate Statistical Theory*, New Jersey: John Wiley & Sons (1982).