Universidad
Carlos III de Madrid
www.uc3m.es

TESIS DOCTORAL

PARAMETER UNCERTAINTY IN PORTFOLIO OPTIMIZATION

Autor:

Alberto Martín Utrera

Director:

Francisco J. Nogales

Co-Director:

Victor DeMiguel

DEPARTAMENTO DE ESTADÍSTICA

Getafe, Julio 2013

Universidad Carlos III de Madrid
www.uc3m.es

TESIS DOCTORAL

PARAMETER UNCERTAINTY IN PORTFOLIO OPTIMIZATION

Autor: *Alberto Martín Utrera*

Director: Francisco J. Nogales Martín

Co-Director: Victor DeMiguel

Firma del Tribunal Calificador:

Firma

Presidente:

Vocal:

Secretario:

Calificación:

Getafe, de de

UNIVERSIDAD CARLOS III DE MADRID

Dissertation

Parameter Uncertainty in Portfolio Optimization

Author: **Alberto Martín-Utrera**

Advisors:
Francisco J. Nogales
Universidad Carlos III de Madrid
Victor DeMiguel
London Business School

“You don’t choose a life, you live one”

Contents

List of Figures

List of Tables

Acknowledgements

I would like to thank my advisors Francisco Javier Nogales and Victor DeMiguel for their help and encouraging advise during my dissertation. I want to stress my gratitude to them for all the academic and personal opportunities they have given me. Javier, I am in debt with you for your calmed, polite, kind and supportive guidance during the last four years. Victor, I will always be grateful to you for letting me work with you at London Business School, widen my ambitions in life, and push me to achieve my dreams. This thesis started as a professional project that turned into something personal and both of you lift me up every time I fell along the way. I will always keep in my heart the memories of this journey and your friendship.

I also want to express my gratitude to all members of the Department of Statistics at University Carlos III of Madrid. Particularly to Gabi, Miguel, Sofi, Nicola, Cris, Henry, Dalia, Ana, Pau, Belen, Carlo, Diego, Jorge, Bernardo, Javier, Sergio, and Esther. Special thanks to my officemate Leo for all the hard moments we shared during the Master and all the nice conversations (sometimes very philosophical) we had in the office. I also want to thank my Phd fellows for all the nice moments we had in Getafe, specially to Jonatan Groba for his priceless friendship and his patience with me every time I had questions about LaTex.

Thanks to Woonam, Iva, Xue, Jonas, Heikki, Yiangos and Isabel, my friends from London Business School. You definitely made my time in London a fantastic experience that I will never forget.

I also want to thank my childhood friends Villa, Mingui, Isra, Charly, Sesmero, Alberto, David and Pau. Words can't express my gratitude to all of you. You are the true engine of my life and you always give me that moment of fun and joy when I mostly need it. I can proudly say that I have the privilege of feeling what true friendship is like. Thank you guys.

Last but not least (and actually first) I want to thank my parents, my brother and Miriam for their support and sacrifice. They represent a clear example of unconditional love. Thanks to my parents for giving me everything they never had and for letting me be the person I want to be. Thanks to my brother and Miriam for always cheering me up and believing in me more than I do. Mum, Dad, brother, and Miriam, you are my family and *everything I do, I do it for you.*[†]

†This is a quotation from the song "(Everything I do) I do it for you" by Canadian signer Bryan Adams. Eight simple words properly put together can say so much about what one feels, and this is what I feel about my family.

Resumen

La modelización de decisiones reales supone la interacción de dos elementos: un problema de optimización y un procedimiento para estimar los parámetros que definen dicho modelo. Cualquier técnica de estimación requiere de la utilización de información muestral disponible, la cual es aleatoriamente dada. Dependiendo de dicha muestra, los estimadores pueden variar ampliamente, y en consecuencia uno puede obtener soluciones muy distintas del modelo. Concretamente, la incertidumbre de los estimadores que definen el modelo resulta en decisiones inciertas.

El análisis del impacto de la incertidumbre de los parámetros en la optimización de carteras es un área muy activo en estadística e investigación operativa. En esta tesis tratamos el impacto de la incertidumbre de los parámetros en la optimización de carteras. En concreto, estudiamos y caracterizamos la pérdida esperada de los inversores que usan información muestral para construir sus carteras óptimas, y además proponemos nuevas técnicas para aliviar dicha incertidumbre.

Primero estudiamos diferentes criterios de calibración para estimadores *shrinkage* en el contexto de la optimización de carteras. En concreto consideramos diferentes métodos de calibración para estimadores shrinkage del vector de medias, la matriz de covarianzas y el vector de pesos. Para cada método de calibración damos expresiones explícitas de la intensidad óptima del shrinkage y además proponemos un nuevo enfoque no-paramétrico para el cálculo de la intensidad de shrinkage de cada criterio de calibración. Finalmente evaluamos el comportamiento de cada método de calibración con datos simulados y empíricos.

En segundo lugar analizamos el impacto de la incertidumbre de los parámetros para un inversor multiperiodo que se enfrenta a costes de transacción. Caracterizamos la pérdida esperada del inversor multiperiodo y encontramos que dicha pérdida es igual al producto de la perdida de un solo periodo y otro término que recoge los efectos multiperiodo en la perdida de utilidad. Además proponemos dos carteras multiperiodo de tipo shrinkage que ayudan a mitigar la incertidumbre de los parámetros. Finalmente analizamos el comportamiento de las carteras multiperiodo que proponemos y encontramos que el inversor puede sufrir grandes pérdidas si ignora los costes de transacción, la incertidumbre de los parámetros o ambos elementos.

Abstract

Modeling every real-world decision involves two elements: an optimization problem and a procedure to estimate the parameters of the model. Any estimation technique requires the utilization of available sample information, which is random. Depending on the given sample, the estimates may vary widely, and in turn, one may obtain very different solutions from the model. Precisely, the uncertainty of the estimates that define the parameters of the model results into uncertain decisions.

Analyzing the impact of parameter uncertainty in optimization models is an active area of study in statistics and operations research. In this dissertation, we address the impact of parameter uncertainty within the context of portfolio optimization. In particular, we study and characterize the expected loss for investors that use sample estimators to construct their optimal portfolios, and we propose several techniques to mitigate the impact of parameter uncertainty.

First, we study different calibration criteria for shrinkage estimators in the context of portfolio optimization. Precisely, we study shrinkage estimators for both the inputs and the output of the portfolio model. In particular, we consider a set of different calibration criteria to construct shrinkage estimators for the vector of means, the covariance matrix, and the vector of portfolio weights. We provide analytical expressions for the optimal shrinkage intensity of each calibration criteria, and in addition, we propose a novel non-parametric approach to compute the optimal shrinkage intensity. We characterize the out-of-sample performance of shrinkage estimators for portfolio selection with simulated and empirical datasets.

Second, we study the impact of parameter uncertainty in multiperiod portfolio selection with transaction costs. We characterize the expected loss of a multiperiod investor, and we find that it is equal to the product between the single-period utility loss and a second term that captures the multiperiod effects on the overall utility loss. In addition, we propose two multiperiod shrinkage portfolios to mitigate the impact of parameter uncertainty. We test the out-of-sample performance of these novel multiperiod shrinkage portfolios with simulated and empirical datasets, and we find that ignoring transaction costs, parameter uncertainty, or both, results into large losses in the investor's performance.

1. Introduction

The basic ideas of asset selection arose in 1952 with Markowitz's paper *Portfolio Selection*, published in *The Journal of Finance*. In his paper, Markowitz introduces a mathematical formulation to accomplish efficient investments based on risk and expected return. Markowitz's work founded the principles of risk diversification:

> *The basic concepts of portfolio theory came to me one afternoon in the library while reading John Burr Williams's Theory of Investment Value. Williams proposed that the value of a stock should equal the present value of its future dividends. Since future dividends are uncertain, I interpreted Williams's proposal to be to value a stock by its expected future dividends. But if the investor were only interested in expected values of securities, he or she would only be interested in the expected value of the portfolio; and to maximize the expected value of a portfolio one need invest only in a single security. This, I knew, was not the way investors did or should act. Investors diversify because they are concerned with risk as well as return. Variance came to mind as a measure of risk. The fact that portfolio variance depended on security covariances added to the plausibility of the approach. Since there were two criteria, risk and return, it was natural to assume that investors selected from the set of Pareto optimal risk-return combinations.*[1]

Markowitz defines the variance of asset returns as a measure of risk to construct an optimization problem that represents the fundamental investor's challenge. In Markowitz's framework, the investor aims to find the optimal combination of assets that minimizes the portfolio variance under a given expected portfolio return, which is known as the mean-variance framework. This problem can be formulated as follows:

$$\min_{w} \quad w' \Sigma w \tag{1.1}$$

$$\text{s.t.} \quad w' \mu \geq \overline{\mu}, \tag{1.2}$$

$$w' \iota = 1, \tag{1.3}$$

where μ is the vector of expected returns, Σ is the covariance matrix of returns, $\overline{\mu}$ is the target expected return, and ι is an appropriate vector of ones. This problem has an equivalent formulation that admits an explicit solution:

$$\max_{w} \quad \mu' w - \frac{\gamma}{2} w' \Sigma w \tag{1.4}$$

$$\text{s.t.} \quad w' \iota = 1, \tag{1.5}$$

[1]Quotation from Markowitz's autobiography in:

http://nobelprize.org/nobel_prizes/economics/laureates/1990/markowitz-autobio.html where γ is the investors risk-aversion parameter. The solution to this problem is obtained directly by applying the *first order conditions* (FOCs):

$$w = \frac{1}{\gamma}\Sigma^{-1}\mu - \frac{\lambda}{\gamma}\Sigma^{-1}\iota, \tag{1.6}$$

where λ is the Lagrange multiplier of constraint (1.5) and is defined as $\lambda = \frac{\mu'\Sigma\iota - \gamma}{\iota'\Sigma^{-1}\iota}$.

Under the mean-variance framework, the investor obtains those portfolios on the so-called efficient frontier, which contains the portfolios that provide an efficient trade-off between risk and return. This is a paradigm of the concept of risk diversification, where one can attain the lowest risk for a given expected return. Through diversification, investors can remove the idiosyncratic risk, or risk inherent to each individual asset, and thus the efficient portfolio is only affected by systemic risk.

1.1. Multiperiod portfolio optimization

The classical mean-variance framework does not account for possible changes in the investment opportunity set. Hence, today's asset allocation decision is based on a static problem where the investor is only concerned about tomorrow's payoff and not about future payoffs. However, investors trade actively and they periodically rebalance their portfolios in accordance with the dynamics of market conditions. To model this realistic situation, one may use dynamic programming. In the dynamic (multiperiod) setting, the investor maximizes her expected utility accumulated along a finite or an infinite investment horizon.

To model an investment strategy, we must define the utility function that the investor wants to maximize. For instance, we can assume that the investor aims to maximize a concave utility defined by the accumulated wealth:

$$W_{t+1} = W_t(w_t'R_{t+1} + r_t^f), \tag{1.7}$$

where W_t is the investor's wealth at time t, w_t is the investor's portfolio at time t, R_t is the vector of stock returns, and r_t^f is the return of the risk-free asset at time t. Expression (1.7) is also known as the budget constraint, and it has to be satisfied along the multiperiod decision problem, which is defined as

$$\begin{aligned} V(\tau, W_t, z_t) &= \max_{\{w_s\}_{s=t}^{t+\tau-1}} E_t[u(W_{t+\tau})] && (1.8)\\ &= \max_{w_t} E_t\left[\max_{\{w_s\}_{s=t+1}^{t+\tau-1}} E_{t+1}[u(W_{t+\tau})]\right] && (1.9)\\ &= \max_{w_t} E_t\left[V\left(\tau-1, W_{t+1} = W_t(w_t'R_{t+1} + r_t^f), z_{t+1}\right)\right], && (1.10)\end{aligned}$$

where τ is the investment horizon, and z_t is a vector of state variables. Equation (1.10) is known as the Bellman equation. In general, this problem can be solved numerically using the system of nonlinear equations obtained from the First Order Conditions (FOCs). For a better understanding, let us consider the particular example of an investor with constant relative risk aversion (CRRA) preferences. In this specific context, the investor's utility is of the form $u(W_{t+\tau}) = W_{t+\tau}^{1-\gamma}/(1-\gamma)$, where γ is the investor's relative risk aversion coefficient. Then, the investment decision problem can be established as follows:

$$V(\tau, W_t, z_t) = \max_{w_t} E_t \left[\max_{\{w_s\}_{s=t+1}^{t+\tau-1}} E_{t+1} \left[\frac{W_{t+\tau}^{1-\gamma}}{1-\gamma} \right] \right] \quad (1.11)$$

$$= \max_{w_t} E_t \left[\max_{\{w_s\}_{s=t+1}^{t+\tau-1}} E_{t+1} \left[\frac{(W_t \prod_{s=t}^{t+\tau-1} (w_s^T R_{s+1} + r_s^f))^{(1-\gamma)}}{1-\gamma} \right] \right] \quad (1.12)$$

$$= \max_{w_t} E_t \left[\frac{W_{t+1}^{(1-\gamma)}}{1-\gamma} \max_{\{w_s\}_{s=t+1}^{t+\tau-1}} E_{t+1} \left[\left(\prod_{s=t+1}^{t+\tau-1} (w_s^T R_{s+1} + r_s^f) \right)^{(1-\gamma)} \right] \right] \quad (1.13)$$

$$= \max_{w_t} E_t \left[u(W_{t+1}) \Psi(\tau - 1, z_{t+1}) \right]. \quad (1.14)$$

Therefore, the value function $V(\tau, W_t, z_t)$ is equal to the product of the investor's utility at $t+1$, defined by the wealth at time $t+1$, and $\Psi(\tau-1, z_{t+1})$ that represents the discounted expected value function with an investment horizon of $\tau - 1$. Note that when $\tau = 1$, we deal with the *myopic* portfolio selection problem.

In this thesis, we mainly focus on a mean-variance investor whose objective is to maximize her expected portfolio return penalized by the portfolio variability. In particular, we study both a myopic investor ($\tau = 1$) and a multiperiod/dynamic investor ($\tau > 1$) in Chapters 2 and 3, respectively.

1.2. Parameter uncertainty

One of the most important challenges in portfolio optimization is the impact of parameter uncertainty in the investor's performance. In general, investment decisions are formulated as optimization problems where the investor maximizes the expected utility, which is defined by a set of parameters that are unknown and the investor has to estimate them. Portfolios constructed with sample estimates may result into suboptimal decisions, and this effect worsens the investor's expected performance.

In the mean-variance framework, parameter uncertainty has an important relevance for two main reasons:

1. The mean-variance framework is very sensitive to changes affecting the inputs of the problem and small distortions may provide extremely different optimal portfolios. Thus, small errors affecting the inputs within the mean-variance framework can give very suboptimal results. Consequently, the mean-variance framework is known as an *error maximizer*; see Michaud (1989).

2. The fragility of the mean-variance framework to changes in the inputs of the model provide very extreme portfolios which may result into prohibitive transaction costs for active managers.

Mean-variance investors that want to mitigate the impact of parameter uncertainty have to take into account the aforementioned problems. We now provide a small review of some of the methods considered in the literature to reduce the drawbacks that arise from the mean-variance model.

1.2.1. Factor models

One possibility to reduce the impact of parameter uncertainty in portfolio optimization is to reduce the dimensionality of the problem and in turn, reduce the number of parameters. We can do that using factor models; see Chan et al. (1999).

When we deal with a portfolio of N assets, there are $\frac{N(N+1)}{2}$ parameters to be estimated only in the covariance matrix. Using a factor model to describe the dynamics of assets returns we can considerably reduce the number of parameters of the covariance matrix. For a general K-factor model, we describe the dynamics of asset returns as:

$$r_{it} = \alpha_i + \beta_i' f_t + \epsilon_{it}, \tag{1.15}$$

where α_i is known as the manager's ability coefficient, β_i is the vector of factor loadings, f_t is the vector of factor realizations and ϵ_{it} is the error term. If the considered factors are distributed as $N(\mu_f, \Sigma_f)$ and the error term follow a $N(0, \Sigma_\epsilon)$, then we obtain that the estimates for the unconditional mean and covariance matrix are:

$$\widehat{\mu} = \widehat{\alpha} + \widehat{B}'\widehat{\mu}_f \tag{1.16}$$

$$\widehat{\Sigma} = \widehat{B}\widehat{\Sigma}_f\widehat{B}' + \widehat{\Sigma}_\epsilon, \tag{1.17}$$

where B is the $N \times K$ matrix of factor loadings. Therefore, the number of parameters in the covariance matrix reduces to $K + N(K+1)$ terms when factors are uncorrelated. Specifically, a 3-factor model only requires to estimate $2,003$ parameters to construct the covariance matrix of a portfolio with $N = 500$ assets, in comparison with the $125,250$ parameters that we estimate using the sample covariance matrix.[2]

1.2.2. Resampling methods

Another method to combat the impact of parameter uncertainty is to use resampling methods. With this technique, the investor considers different scenarios and the general solution is obtained by averaging the solutions for each of the considered scenarios. This is an alternative technique popularized by Michaud (1998) that tries to provide more diversified portfolios, unlike the original mean-variance framework which gives extreme portfolio weights.

Resampling the original dataset, we provide an approximation of the different possible scenarios. In particular, we can use bootstrap techniques to construct the different scenarios: (μ_1, Σ_1), …,(μ_B, Σ_B), where B corresponds with the number of all bootstrap estimates for μ and Σ. Then, we construct our mean-variance portfolio by averaging all the optimal portfolios constructed for all different B bootstrap scenarios:

$$\overline{w}_m^{\text{resampled}} = \frac{1}{B}\sum_{b=1}^{B} w_b, \tag{1.18}$$

where w_b is the optimal mean-variance portfolio in (1.6) for the bth bootstrap scenario.

[2]If factors are correlated, the number of parameter that we estimate to construct the covariance matrix are $N(K+1) + K(K+1)/2$, where K is the number of factors.

1.2.3. Robust portfolio selection

Under this approach, μ and Σ belong to uncertainty sets and the investor maximizes her utility for the *worst-case* scenario within the uncertainty sets.[3] For expository reason, we present the simplest case where the mean is defined within an uncertainty set and the covariance matrix is known. In particular, we define the uncertainty set for the mean as:

$$\Omega = \{\mu : (\mu - \widehat{\mu})'\Sigma^{-1}(\mu - \widehat{\mu}) \leq k^2\}, \tag{1.19}$$

where $\widehat{\mu}$ is the sample mean and k determines the size of the uncertainty set. Therefore, the robust formulation for this problem takes the form:

$$\max_{w} \quad \min_{\mu \in \Omega} \quad \left\{\widehat{\mu}'w - \frac{\gamma}{2}w'\Sigma w\right\} \tag{1.20}$$

$$s.t. \quad w'\iota = 1. \tag{1.21}$$

We can provide an alternative and tractable formulation of problem (1.20)-(1.21) in order to obtain the optimal solution of the robust formulation. This is as follows; see Garlappi et al. (2007):

$$\max_{w} \quad \widehat{\mu}^T w - \frac{\gamma}{2}w^T\Sigma w - k\|\Sigma^{-1/2}w\| \tag{1.22}$$

$$\text{s.t.} \quad w^T \iota = 1. \tag{1.23}$$

This tractable formulation takes the form of a Second Order Cone Programming (SOCP) problem, which can be solved efficiently. This problem provides a more conservative solution compared with the classical mean-variance approach because it gives a higher weight to the portfolio variability with the new term $-k\|\Sigma^{-1/2}w\|$. In general, robust formulations give more conservative solutions.

1.2.4. The Black-Litterman model

The Black-Litterman (BL) model was first published by Fisher Black and Robert Litterman in an internal paper of Goldman Sachs in 1990. It was subsequently published in the Journal of Fixed Income as the article *Asset Allocation: Combining Investor Views with Market Equilibrium.* The Black-Litterman model is a Bayesian model applied to portfolio selection which takes into consideration the uncertainty in the vector of expected returns. This is a flexible model that incorporates the views of investors into the portfolio decision problem.

In the Black-Litterman model, and according with the CAPM model, the prior distribution for the vector of means is defined as follows:

$$P(\mu) \sim N(\Pi, \phi\Sigma), \tag{1.24}$$

where Π is the vector of mean excess returns in equilibrium, Σ is the covariance matrix of asset returns and ϕ is a parameter that determines the uncertainty on μ.

[3]See Goldfarb and Iyengar (2003) for an extensive application of this methodology in portfolio optimization.

The views of the investor are incorporated in the conditional distribution as follows:

$$P(X|\mu) \sim N(V^{-1}Q, [V'\Omega V]^{-1}), \tag{1.25}$$

where X is the sample of investor's views, V is the matrix of the asset weights within each view, Q is the matrix of returns for each view and Ω is the covariance matrix for the views of the investor.

From (1.24) and (1.25), we can construct the posterior distribution of returns using the Bayes' rule. Accordingly, Black and Litterman define a normal posterior distribution which allows to incorporate the views of the investors in the mean-variance model.

1.2.5. Minimum-variance

From Merton (1980), it is well known that estimating the vector of means is more difficult than estimating the covariance matrix. In particular, Merton (1980) shows that the variation in the realized market returns is much larger than the variation in the variance rate. That is why disregarding the vector of means helps mitigate the impact of parameter uncertainty on the portfolio model. In turn, one can solve the minimum-variance portfolio problem, which is as follows:

$$\min_{w} \quad w'\widehat{\Sigma}w \tag{1.26}$$

$$\text{s.t.} \quad w'\,\iota = 1. \tag{1.27}$$

This model can be solved explicitly, which results into the following optimal portfolio:

$$w = \frac{\Sigma^{-1}\iota}{\iota'\Sigma^{-1}\iota}. \tag{1.28}$$

Although the error affecting the covariance matrix is lower than the vector of means, this element, however, also suffers from estimation risk and this may provide large losses. In the following section, we give details of a statistical technique that reduces the estimation error on Σ.

1.2.6. Shrinkage estimators

There is a large literature of shrinkage estimators within the context of portfolio selection. Among others, Ledoit and Wolf (2003) and Ledoit and Wolf (2004a) develop shrinkage estimators for the covariance matrix and they give empirical evidences of their good performance in portfolio selection. In general, they propose estimators for the covariance matrix that result from an optimal combination between the sample covariance matrix and a target covariance matrix:

$$\widehat{\Sigma}_{\text{Shrink}} = \widehat{\delta}\,F + (1-\widehat{\delta})\widehat{\Sigma}, \tag{1.29}$$

where F represents a target matrix, and $\widehat{\delta}$ is a value between 0 and 1 that corresponds with the shrinkage intensity.

The basic idea of this methodology is that those estimated coefficients in the sample covariance matrix that are extremely high or extremely low tend to contain large estimation error. Thus, it is optimal to push those extreme values to most centered values in order to reduce estimation error.

As a result, parameter δ is the key element and it determines the performance of these estimators. In Ledoit and Wolf (2003) and Ledoit and Wolf (2004a), the authors propose a quadratic loss function to compute parameter δ. In particular, the aim is to find the optimal value of δ that minimizes the following loss function:

$$\min_{\delta} E\left(\|\delta F + (1-\delta)\widehat{\Sigma} - \Sigma\|_F \right), \tag{1.30}$$

where Σ is the population covariance matrix.

1.2.7. Constrained minimum-variance

Imposing short-selling constraints is another mechanism to diminish the estimation error into the variance-covariance matrix. Jagannathan and Ma (2003) show that constraining short-selling demands has an effect on the covariance matrix similar to that of shrinkage estimators. The optimization problem with short-selling constraints and upper bounds is:

$$\min_{w} \quad w'\widehat{\Sigma}w \tag{1.31}$$
$$\text{s.t.} \quad w'\iota = 1 \tag{1.32}$$
$$w_i \geq 0, \; i = 1, \dots, N \tag{1.33}$$
$$w_i \leq \overline{w}, \; i = 1, \dots, N, \tag{1.34}$$

where $\overline{w}$ is the upper bound of the portfolio weights.

Jagannathan and Ma (2003) show that the effect of portfolio constraints on the sample covariance matrix is equivalent to estimate the sample covariance matrix as:

$$\widetilde{\Sigma} = \widehat{\Sigma} + (\delta\iota' + \iota\delta') - (\lambda\iota' + \iota\lambda'), \tag{1.35}$$

where δ is the vector of Lagrange multipliers for the upper bound constraints and λ is the vector of Lagrange multipliers for the short-selling constraints. Each Lagrange multiplier is positive when the corresponding constraint is *active* and zero otherwise. Expression (1.35) basically shows that when the short-sell constraint of stock i is active ($\lambda_i \geq 0$) and the upper limit constraint is not active ($\delta = 0$), the covariance is reduced to $\widetilde{\sigma}_{i,j} = \widehat{\sigma}_{i,j} - \lambda_i - \lambda_j$. Overall, the covariance matrix of a constrained minimum-variance problem acts as a shrinkage estimator.

1.2.8. Norm-constrained model

DeMiguel et al. (2009) provide a general framework where portfolio weights are *norm-constrained*. Accordingly, the norm of portfolio weights must be smaller than a given threshold. This framework includes in formulation (1.26)-(1.27) the additional constraint of the norm of the portfolio weights:

$$\min_{w} \quad w^T\widehat{\Sigma}w \tag{1.36}$$
$$\text{s.t.} \quad w^T\,\iota = 1 \tag{1.37}$$
$$\|w\|_N \leq \delta, \tag{1.38}$$

where $\|w\|_N$ is any norm of portfolio weights. DeMiguel et al. (2009) show that when the norm-constraint is the 1-norm,$\|w\|_1 = \sum_{i=1}^{N} |w_i|$, with $\delta = 1$ the model gives the same result as the short-constrained minimum-variance. Moreover, they show that the A-norm minimum-variance problem gives the same portfolio as the unconstrained minimum-variance problem with the shrinkage estimator of Ledoit and Wolf when with $\widehat{\Sigma}_{LW} = (1-\nu)\widehat{\Sigma} + \nu A$.[4] Finally, they show that when $A = I$, the A-norm is simply the 2-norm, $\|w\|_2 = \sum_{i=1}^{N} w_i^2$. In that specification, if the threshold parameter δ is equal to $\frac{1}{N}$, the solution of the problem is exactly the equally weighted portfolio.

1.2.9. The naive strategy

The naive strategy does not use any statistical tool to construct the investor's portfolio and it simply allocates the same budget proportion to every risky asset. Although this is not a very sophisticated technique, the naive portfolio has a good out-of-sample performance because is not affected by estimation risk. In turn, DeMiguel et al. (2009) conclude that the gain associated to optimal and well diversified portfolios is smaller than the loss coming from estimation error.

1.3. Our contribution

In thesis, we deal with the impact of parameter uncertainty in portfolio optimization. In a broad sense, we contribute to the literature in two main areas: First, we focus on the single-period portfolio problem and we study a wide variety of shrinkage methods for portfolio selection. We propose new different calibration criteria and we extensively test their out-of-sample performance in an empirical application. Second, we deal with the impact of parameter uncertainty for multiperiod investor that also suffers from transaction costs. We analytically characterize the investor's expected loss and we propose two multiperiod shrinkage portfolios that considerably improve the investor's out-of-sample performance.

Specifically, in Chapter 2 we carry out a comprehensive investigation of shrinkage estimators for asset allocation, and we find that *size matters*—the shrinkage intensity plays a significant role in the performance of the resulting estimated optimal portfolios. We study both portfolios computed from shrinkage estimators of the moments of asset returns (*shrinkage moments*), as well as *shrinkage portfolios* obtained by shrinking the portfolio weights directly. We make several contributions in this field. First, we propose two novel calibration criteria for the vector of means and the inverse covariance matrix. Second, for the covariance matrix we propose a novel calibration criterion that takes the condition number optimally into account. Third, for shrinkage portfolios we study two novel calibration criteria. Fourth, we propose a simple multivariate smoothed bootstrap approach to construct the optimal shrinkage intensity. Finally, we carry out an extensive out-of-sample analysis with simulated and empirical datasets, and we characterize the performance of the different shrinkage estimators for portfolio selection.

In Chapter 3, we study the impact of parameter uncertainty in multiperiod portfolio selection with trading costs. We analytically characterize the expected loss of a multiperiod investor, and we find that it is equal to the product of two terms. The first term

[4] The A-norm is defined as $\|w\|_A = (w'Aw)^{1/2}$ with $A \in \mathbb{R}^{N\times N}$ is a positive matrix

corresponds with the single-period utility loss in the absence of transaction costs, as characterized by Kan and Zhou (2007), whereas the second term captures the multiperiod effects on the overall utility loss. To mitigate the impact of parameter uncertainty, we propose two multiperiod shrinkage portfolios. The first multiperiod shrinkage portfolio combines the Markowitz portfolio with a target portfolio. This method diversifies the effects of parameter uncertainty and reduces the risk of taking inefficient positions. The second multiperiod portfolio shrinks the investor's trading rate. This novel technique smooths the investor trading activity and it helps to reduce the impact of parameter uncertainty. Finally, we test the out-of-sample performance of our considered portfolio strategies with simulated and empirical datasets, and we find that ignoring transaction costs, parameter uncertainty, or both, results into large losses in the investor's performance.

We conclude and summarize the main findings of this thesis in Chapter 4, and we also give a short description of possible research lines in Chapter 4.2.

2. Size Matters: Optimal Calibration of Shrinkage Estimators for Portfolio Selection

2.1. Overview

The classical mean-variance framework for portfolio selection proposed by Markowitz (1952) formalizes the concept of investment diversification, and it is widely used nowadays in the investment industry. To compute mean-variance portfolios, one needs to estimate the mean and covariance matrix of asset returns. One possibility is to replace these quantities with their sample estimators, but these are obtained from historical return data and contain substantial estimation error. As a result, mean-variance portfolios computed from sample estimators perform poorly out of sample; see, for instance, Jobson and Korkie (1981); Best and Grauer (1991); Broadie (1993); Britten-Jones (1999); DeMiguel, Garlappi, and Uppal (2009).

One of the most popular approaches to combat the impact of estimation error in portfolio selection is to use shrinkage estimators, which are obtained by "shrinking" the sample estimator towards a target estimator.[1] The advantage is that while the shrinkage target is usually biased, it also contains less variance than the sample estimator. Thus it is possible to show under general conditions that there exists a shrinkage *intensity* for which the resulting *shrinkage* estimator contains less estimation error than the original sample estimator; see James and Stein (1961). The key then is to characterize the optimal trade-off between the sample estimator (low bias), and the target (low variance). In other words, shrinkage estimators can help reduce estimation error, but the shrinkage intensity (*size*) matters.

In this chapter, we make an extensive investigation of shrinkage estimators for portfolio selection. We study both portfolios computed from shrinkage estimators of the moments of asset returns (*shrinkage moments*), as well as *shrinkage portfolios* obtained by shrinking directly the portfolio weights computed from the original (un-shrunk) sample moments.

Constructing shrinkage estimators is a three-step procedure. First, define the shrinkage target. Second, choose the calibration criterion that determines the shrinkage intensity. Third, use the available data to estimate the shrinkage intensity that optimizes the calibration criteria. Our work contributes mainly to the last two steps by proposing new calibration criteria, and providing parametric and nonparametric approaches to compute

[1] Other approaches proposed to combat estimation error in portfolio selection include: Bayesian methods (Barry (1974), Bawa et al. (1979)), Bayesian methods with priors obtained from asset pricing models (MacKinlay and Pastor (2000), Pastor (2000), Pastor and Stambaugh (2000)), robust optimization methods (Cornuejols and Tutuncu (2007), Goldfarb and Iyengar (2003), Garlappi et al. (2007), Rustem et al. (2000), Tutuncu and Koenig (2004)), Bayesian robust optimization (Wang (2005)), robust estimation methods (DeMiguel and Nogales (2009)), and imposing constraints (Best and Grauer (1992), Jagannathan and Ma (2003), and DeMiguel et al. (2009)).

the shrinkage intensity. The shrinkage targets we consider are in general similar to those considered in the existent literature.

We consider three shrinkage estimators of the moments of asset returns. First, we consider a shrinkage estimator of the vector of means similar to those considered before by Jorion (1986) or Frost and Savarino (1986). Unlike these authors, however, we define our estimator a priori as a convex combination of the sample mean and a target element, and we calibrate the shrinkage intensity to minimize the expected quadratic loss—a criterion that distinguishes our work from that of the aforementioned papers. We provide a closed-form expression for the optimal shrinkage intensity under the assumption that returns are independent and identically distributed (iid), but without imposing any further assumptions on the return distribution. Second, we consider the shrinkage covariance matrix proposed by Ledoit and Wolf (2004b), and we implement the same calibration criterion, the expected quadratic loss. Unlike Ledoit and Wolf (2004b), however, we provide a closed-form expression of the optimal shrinkage intensity *for finite samples* by assuming that returns are iid *normal*. Third, we consider a shrinkage estimator of the inverse covariance matrix that is a convex combination of the inverse of the sample covariance matrix and the identity matrix. This estimator is similar to those considered by Frahm and Memmel (2010) and Kourtis et al. (2012), but our contribution is to consider a different calibration criterion for the shrinkage intensity: the expected quadratic loss. Moreover, under iid normal returns, we provide a closed-form expression of the true optimal shrinkage intensity that minimizes the expected quadratic loss. Finally, we propose a new calibration criterion for the shrinkage covariance matrix that takes into account not only the expected quadratic loss but also its condition number. The condition number gives a bound for the sensitivity of the computed portfolio weights to estimation errors in the mean and covariance matrix of asset returns, and thus calibrating the shrinkage covariance matrix so that its condition number is relatively small helps to reduce the impact of estimation error in portfolio selection. Indeed, our experiments with simulated and empirical data demonstrate the advantages of using this criterion for the construction of minimum-variance portfolios.

We investigate three different shrinkage portfolios. The first is obtained by shrinking the sample mean-variance portfolio towards the sample minimum-variance portfolio and it is closely related to the three-fund portfolio of Kan and Zhou (2007); the second is obtained by shrinking the sample mean-variance portfolio towards the equally-weighted portfolio as in Tu and Zhou (2011); and the third is obtained by shrinking the sample minimum-variance portfolio towards the equally-weighted portfolio, similar to DeMiguel, Garlappi, and Uppal (2009). We contribute to the literature by considering, in addition to the utility and variance criteria, two novel calibration criteria: the expected quadratic loss minimization criterion, and the Sharpe ratio maximization criterion. We study the expected quadratic loss criterion because of its good performance in the context of shrinkage covariance matrices (see Ledoit and Wolf (2004a)); and we consider the Sharpe ratio criterion because it is a particular case of the expected utility criterion and it is a relevant performance measure for investors.

For both types of shrinkage estimators, moments and portfolio weights, we propose a multivariate nonparametric smoothed bootstrap approach to estimate the optimal shrinkage intensity. This approach does not impose any assumption on the distribution of asset returns. To the best of our knowledge, this is the first work to consider such a nonparametric approach for shrinkage estimators within the context of portfolio optimization.

Finally, we evaluate the out-of-sample performance of the portfolios obtained from shrinkage moments, as well as that of the shrinkage portfolios on the six empirical datasets listed in Table 2.1. For portfolios computed from shrinkage moments, we identify two main findings. First, the shrinkage estimator of the vector of means calibrated with our proposed criterion improves the out-of-sample performance of the resulting mean-variance portfolios. Second, taking the condition number of the estimated covariance matrix into account improves the quality of its shrinkage estimators. For shrinkage portfolios we identify two main findings. First, we find that for those shrinkage portfolios that make use of the sample mean, the best calibration criterion is the portfolio variance minimization criterion. Second, for shrinkage portfolios that ignore the sample mean, the best calibration criterion is to minimize the expected quadratic loss. Finally, for both shrinkage moments and shrinkage portfolios, we find that the nonparametric bootstrap approach to estimate the optimal shrinkage intensity tends to work better than the parametric approach based on normality.

Summarizing, we contribute to the literature of shrinkage estimators for portfolio selection in the following aspects: first, we propose new calibration criteria for shrinkage estimators of moments of asset returns. Second, we consider new calibration criteria for shrinkage portfolios. Concretely, we consider a expected quadratic loss minimization criterion, as well as a Sharpe ratio maximization criterion. Third, we study a multivariate nonparametric approach to compute the optimal shrinkage intensity when returns are iid. Finally, we carry out a comprehensive empirical investigation of shrinkage estimators for portfolio selection on six empirical datasets.

The chapter is organized as follows. Section 2.2 introduces all the considered shrinkage estimators for portfolio selection. Section 2.3 characterizes the optimal shrinkage intensities when asset returns are iid normal. Section 2.4 proposes a smoothed bootstrap approach to approximate the optimal shrinkage intensities when asset returns are just iid. Section 2.5 gives the results of the simulation experiment, and Section 2.6 compares the performance of the different shrinkage estimators on six empirical datasets. Section 2.7 provides a summary of the chapter.

2.2. Shrinkage estimators for portfolio selection

In the classical mean-variance analysis proposed by Markowitz (1952) the investor aims to maximize her risk-adjusted portfolio return. To formalize the investment problem, one has to define the dynamics of asset returns R_t. It is common in the literature to assume that asset returns, in excess of the risk-free asset, are independent and identically distributed (iid) with vector of means μ and covariance matrix Σ. In this context, a mean-variance investor who wants to invest in a set of N available risky assets solves the following optimization problem:

$$\max_{w} \quad w'\mu - \frac{\gamma}{2} w'\Sigma w, \tag{2.1}$$

where γ is the investor's absolute risk aversion, and w is the vector of portfolio weights. The above formulation can be solved in closed-form, and it takes the expression:

$$w = \frac{1}{\gamma}\Sigma^{-1}\mu. \tag{2.2}$$

An important challenge is that it is well known (Michaud (1989)) that the portfolio given by (2.2) is very sensitive to even small estimation errors in μ and Σ. Shrinkage estimators are one of the most effective approaches to mitigate the impact of estimation error in portfolio optimization. One can apply shrinkage estimators to the estimates of the inputs μ and Σ, and also in the output of the problem w. We consider shrinkage estimators for both, the inputs and the output of the investor's portfolio problem.

2.2.1. Shrinkage estimators of moments

We study shrinkage estimators for the *sample* estimates of the vector of means, the covariance matrix, and the inverse covariance matrix. We consider the unbiased sample vector of means, which is defined as $\mu_{sp} = (1/T)\sum_{t=1}^{T} R_t$, and the unbiased *sample* covariance matrix, which is $\Sigma_{sp} = (1/(T-1))\sum_{t=1}^{T}(R_t - \mu_{sp})(R_t - \mu_{sp})'$, where T is the sample size and subindex sp stands for sample estimator. All the considered shrinkage estimators are defined as a convex combination between the sample estimator and a scaled shrinkage target:

$$\mu_{sh} = (1-\alpha)\mu_{sp} + \alpha\nu\mu_{tg}, \tag{2.3}$$

$$\Sigma_{sh} = (1-\alpha)\Sigma_{sp} + \alpha\nu\Sigma_{tg}, \tag{2.4}$$

$$\Sigma_{sh}^{-1} = (1-\alpha)\Sigma_{sp}^{-1} + \alpha\nu\Sigma_{tg}^{-1}, \tag{2.5}$$

where α is the shrinkage intensity and ν is a scaling parameter that we adjust to minimize the bias of the shrinkage target. The shrinkage intensity α determines the "strength" with which the sample estimator is shrunk towards the scaled shrinkage target, and it takes values between zero and one. When the "strength" is one, the shrinkage estimator equals the scaled shrinkage target, and when α is zero, the shrinkage estimator equals the sample estimator.

We introduce the scaling parameter ν for two reasons. First, the scaling parameter yields a more general type of combination between the sample estimator and the target than just a convex combination. Second, we adjust the scaling parameter to reduce the bias of the shrinkage target. In the case where the calibration criterion is the quadratic loss, this results in a higher optimal shrinkage intensity α than that for the case without the scaling parameter. This is likely to result in more stable estimators that are more resilient to estimation error.

Shrinkage estimator of mean returns

Several Bayesian approaches proposed in the existent literature provide estimators of mean returns that can be interpreted as shrinkage estimators. Frost and Savarino (1986) assume an informative Normal-Wishart conjugate prior where all stocks have the same mean, variance and covariances. The predictive mean turns out to be a weighted average of the sample mean and a prior mean, defined as the historical average return for all stocks.[2] Jorion (1986) estimates the vector of means by integrating a predictive density

[2]For computational convenience, we do not consider this shrinkage estimator in the analysis as a benchmark. This shrinkage vector of means requires the definition of a parameter that determines the strength of belief in the prior mean. Frost and Savarino (1986) propose to estimate this parameter in an Empirical-Bayes fashion by maximizing the likelihood of the prior distribution. Since we do not have closed-form expression for the shrinkage intensity, we do not consider it as a benchmark, but rather we consider the shrinkage estimator proposed by Jorion (1986), which offers a closed-form

function defined by an exponential prior which is only specified for the vector of means. The resulting estimator is defined as a weighted average of the sample mean μ_{sp} and the minimum-variance portfolio mean return.

We consider a shrinkage estimator of means similar to those proposed by Frost and Savarino (1986) and Jorion (1986), but we propose a different calibration criterion. Concretely, we consider a shrinkage estimator that is a weighted average of the sample mean and the scaled shrinkage target $\nu\mu_{tg} = \nu\iota$, where ι is the vector of ones and ν is a scaling factor. We then calibrate ν to minimize the bias of the shrinkage target; that is, $\nu_\mu = \text{argmin}_\nu \|\nu\iota - \mu\|_2^2 = (1/N)\sum_{i=1}^N \mu_i = \overline{\mu}$, and we choose the shrinkage intensity α to minimize the expected quadratic loss of the shrinkage estimator:

$$\min_{\alpha} \quad E\left[\|\mu_{sh} - \mu\|_2^2\right] \tag{2.6}$$

where $\|x\|_2^2 = \sum_{i=1}^N x_i^2$.

Our motivation to consider the expected quadratic loss minimization criterion is that it has been shown that works well in the context of shrinkage estimators of the covariance matrix for portfolio optimization (Ledoit and Wolf (2004a)), and hence we are interested in studying whether the performance of this criterion for the vector of means is also good in the context of portfolio optimization. Another advantage of using this criterion is that we are able to provide closed-form expressions for the optimal shrinkage intensity without imposing any assumption on the distribution of asset returns, other that they are iid.

The following proposition gives the true optimal value of the shrinkage intensity α.

Proposition 1. *Assuming asset returns are iid, the shrinkage intensity α that minimizes the expected quadratic loss is:*

$$\alpha_\mu = \frac{E\left(\|\mu_{sp} - \mu\|_2^2\right)}{E\left(\|\mu_{sp} - \mu\|_2^2\right) + \|\nu_\mu\iota - \mu\|_2^2} = \frac{(N/T)\overline{\sigma^2}}{(N/T)\overline{\sigma^2} + \|\nu_\mu\iota - \mu\|_2^2}, \tag{2.7}$$

where $\overline{\sigma^2} = trace\,(\Sigma)\,/N$.

Note that the true optimal shrinkage intensity α is defined by the relative expected loss of the sample vector of means with respect to the total expected loss, defined by the expected loss of the sample vector of means plus the loss of the scaled vector of ones. We observe that the shrinkage intensity increases with the number of assets N, and decreases with the number of observations.

The main difference between the shrinkage estimator we consider and those proposed by Frost and Savarino (1986) and Jorion (1986) is that we use the expected quadratic loss as the calibration criterion. As a result, unlike the Bayes-Stein shrinkage intensity of Jorion (1986), our shrinkage intensity does not depend on the inverse covariance matrix. Accordingly, our optimal shrinkage intensity is easier to compute, particularly when there is a large number of assets and thus constructing the inverse covariance matrix may be computationally expensive.

Shrinkage estimator of the covariance matrix

We consider the shrinkage estimator of the covariance matrix defined by Ledoit and Wolf (2004b), who propose shrinking the sample covariance matrix towards a scaled identity

---

expression of the shrinkage intensity.

matrix; i.e. $\Sigma_{tg} = I$. These authors choose the scaling factor ν to minimize the bias of the shrinkage target I; that is, $\nu_\Sigma = \text{argmin}_\nu \|\nu I - \Sigma\|_F^2 = (1/N)\sum_{i=1}^N \sigma_i^2 = \overline{\sigma^2}$. Under the assumption of iid observations, they propose choosing the shrinkage intensity α to minimize the expected quadratic loss $E\left[\|\Sigma_{sh} - \Sigma\|_F^2\right]$, where $\|X\|_F^2 = \text{trace}(X'X)$. The resulting optimal shrinkage intensity is

$$\alpha_\Sigma = \frac{E\left(\|\Sigma_{sp} - \Sigma\|_F^2\right)}{E\left(\|\Sigma_{sp} - \Sigma\|_F^2\right) + \|\nu_\Sigma I - \Sigma\|_F^2}. \tag{2.8}$$

Similarly to the shrinkage estimator for the vector of means, the shrinkage intensity for the covariance matrix is the relative expected loss of the sample covariance matrix with respect to the expected loss of the sample covariance matrix and the scaled identity matrix. Ledoit and Wolf (2004b) obtain an estimator of α by giving consistent estimators of $E(\|\Sigma_{sp} - \Sigma\|_F^2)$, $\|\nu_\Sigma I - \Sigma\|_F^2$ and ν_Σ. We give on the other hand a closed-form expression of $E(\|\Sigma_{sp} - \Sigma\|_F^2)$ when asset returns are iid normal; see Section 2.3.1. This allows us to better understand the impact of estimation error as a function of the number of assets and observations. Furthermore, in Section 2.4 we propose an alternative nonparametric bootstrap procedure to estimate α for the case of iid returns.

Shrinkage estimator of the inverse covariance matrix

The inverse covariance matrix is a key element to compute mean-variance and minimum-variance portfolios. This is particularly of interest when the number of observations T is not very large relative to the number of assets N, a common situation in portfolio selection. In this case, the covariance matrix is nearly singular and estimation error explodes when we invert it to construct the optimal portfolio weights.

We study a shrinkage estimator of the inverse covariance matrix where the target is the identity matrix (that is, $\Sigma_{tg}^{-1} = I$) and we choose the scaling factor ν to minimizes the bias of the shrinkage target; that is, $\nu_{\Sigma^{-1}} = \text{argmin}_\nu \|\nu I - \Sigma^{-1}\|_F^2 = (1/N)\sum_{i=1}^N \sigma_i^{-2} = \overline{\sigma^{-2}}$. Then, we select the shrinkage intensity α that minimizes the expected quadratic loss $E(\left\|\Sigma_{sh}^{-1} - \Sigma^{-1}\right\|_F^2)$.

Similarly, Frahm and Memmel (2010) and Kourtis et al. (2012) also investigate shrinkage estimators for the inverse covariance matrix. The main difference with our approach is that while Frahm and Memmel (2010) and Kourtis et al. (2012) calibrate their estimators to minimize the out-of-sample portfolio variance, we calibrate our estimator to minimize the expected quadratic loss of the inverse covariance matrix. We select this calibration criterion because Ledoit and Wolf (2004b) show that it results in good performance within the context of the covariance matrix.

From the first-order optimality conditions, we obtain that the optimal shrinkage intensity is

$$\alpha_{\Sigma^{-1}} = \frac{E\left(\left\|\Sigma_{sp}^{-1} - \Sigma^{-1}\right\|_F^2\right) - E\left(< \Sigma_{sp}^{-1} - \Sigma^{-1}, \nu_{\Sigma^{-1}} I - \Sigma^{-1} >\right)}{E\left(\left\|\Sigma_{sp}^{-1} - \Sigma^{-1}\right\|_F^2\right) + \|\nu_{\Sigma^{-1}} I - \Sigma^{-1}\|_F^2 - 2E\left(< \Sigma_{sp}^{-1} - \Sigma^{-1}, \nu_{\Sigma^{-1}} I - \Sigma^{-1} >\right)}, \tag{2.9}$$

where $< A, B >= \text{trace}(A'B)$. Note that the optimal shrinkage intensity is given by the relative expected loss of the inverse of the sample covariance matrix with respect to the expected loss of the inverse of the sample covariance matrix and the scaled identity matrix. In this case, the expected losses are smoothed by an element proportional to the bias of the inverse of the sample covariance matrix. Later, we give closed-form expressions for the expectations in (2.9) under the assumption of iid normal returns, and we also study a nonparametric approach to estimate these expectations assuming just iid returns.

Note that this shrinkage estimator may be very conservative because it is obtained by first inverting the sample covariance matrix, and then shrinking it towards the scaled identity matrix. If the sample covariance matrix is nearly singular, small estimation errors affecting the sample covariance matrix become very large errors in the inverse covariance matrix and this, in turn, results in very large shrinkage intensities (i.e. $\alpha \approx 1$). In this situation, the shrinkage inverse covariance matrix might not capture valuable information about the variances and covariances of asset returns. To address this problem, we propose an alternative calibration criterion in the following section.

Shrinkage estimator of the covariance matrix considering the condition number

We propose an alternative calibration criterion for the covariance matrix that accounts for both the expected quadratic loss and the condition number of the shrinkage covariance matrix, which measures the impact of estimation error on the portfolio weights.[3]

To measure the expected quadratic loss we use the *relative improvement in average loss* (RIAL); see Ledoit and Wolf (2004b):

$$RIAL\left(\Sigma_{sh}\right) = \frac{E\left(\|\Sigma_{sp} - \Sigma\|_F^2\right) - E\left(\|\Sigma_{sh} - \Sigma\|_F^2\right)}{E\left(\|\Sigma_{sp} - \Sigma\|_F^2\right)}. \tag{2.10}$$

The RIAL is bounded above by one, and unbounded below. The maximum value is attained when the expected quadratic loss of the shrinkage estimator Σ_{sh} is negligible relative to the expected quadratic loss of the sample covariance matrix Σ_{sp}. The advantage of using RIAL with respect to using plain expected quadratic loss is that the RIAL is bounded above by one and thus it is easy to compare the RIAL and the condition number of the shrinkage covariance matrix. Note that to characterize the $RIAL\left(\Sigma_{sh}\right)$ it is enough to characterize the expectation $E\left(\|\Sigma_{sp} - \Sigma\|_F^2\right)$. In Section 2.3.1, we give a closed-form expression for this expectation, whereas in Section 2.4 we provide a nonparametric procedure to approximate it.

On the other hand, the condition number of the shrinkage covariance matrix Σ_{sh} is:

$$\delta_{\Sigma_{sh}} = \frac{(1-\alpha)\lambda_{\max} + \alpha\nu_\Sigma}{(1-\alpha)\lambda_{\min} + \alpha\nu_\Sigma}, \tag{2.11}$$

where $\lambda_{\max}$ and $\lambda_{\min}$ are the maximum and minimum eigenvalues of the sample covariance matrix, respectively.[4] The smallest (and thus best) condition number is one, which is attained when α is one. In that case, the shrinkage covariance matrix coincides with the scaled identity matrix.

[3]The condition number is a measure of the matrix singularity, and it provides a bound on the accuracy of the computed solution to a linear system. Mean-variance and minimum-variance portfolios can be interpreted as the solutions of a linear system and this is why the condition number of the estimated covariance matrix matters on the investor's portfolio. Some approaches have been already proposed to deal with this problem by shrinking the eigenvalues of the sample covariance matrix (see Stein (1975), Dey and Srinivasan (1985), Zumbach (2009)).

[4]See Ledoit and Wolf (2004b), equation (13), for the expression of the eigenvalues of the shrinkage covariance matrix. We use that equation to obtain the expression for the condition number of the shrinkage covariance matrix.

Therefore, we propose the following problem to find an optimal shrinkage intensity that accounts both for the expected quadratic loss and the condition number of the shrinkage covariance matrix:

$$\alpha = \text{argmin} \left\{ \delta_{\Sigma_{sh}} - \phi RIAL \left(\Sigma_{sh} \right) \right\}, \tag{2.12}$$

where ϕ is a tuning parameter that controls for the trade-off between the $RIAL$ and the condition number. While the expected quadratic loss and the condition number reflect different properties of the covariance matrix, it is interesting to consider both of them in the calibration procedure. On the one hand, we want to compute portfolios using an estimator of the covariance matrix that is on expectation as close as possible to the true covariance matrix. This can be achieved by minimizing the expected quadratic loss of the estimated covariance matrix. On the other hand, we also want to compute portfolios that are not too sensitive to small changes in the estimated covariance matrix. This can be done by accounting for the condition number of the covariance matrix. These two criteria combined (RIAL and condition number) result in stable portfolios with low expected quadratic loss.

This is a very flexible calibration criterion. In particular, if parameter $\phi = 0$, the objective is to minimize the condition number of Σ_{sh}. In that case, the optimal shrinkage intensity would be one, since that value minimizes the condition number of Σ_{sh}. On the other hand, the larger the value of ϕ, the more important the $RIAL$ is. Then, if $\phi \to \infty$, the above formulation would be equivalent to minimize the expected quadratic loss of the shrinkage matrix.

The parameter ϕ must be exogenously specified. In our empirical analysis, we set ϕ as the value that minimizes the portfolio variance, i.e. $\phi = \arg\min_{\phi} \sigma^2_{\phi}$, where σ^2_{ϕ} is the portfolio variance of the minimum variance portfolio formed with the shrinkage covariance matrix Σ_{sh}, calibrated by criterion (B.2). To compute the portfolio variance, we use the nonparametric technique known as *cross-validation* (see Efron and Gong (1983)).[5] Since problem (B.2) is a highly nonlinear optimization problem, it is difficult to obtain a closed-form solution. Instead, we solve the problem numerically.

2.2.2. Shrinkage estimators of portfolio weights

We now focus on shrinkage portfolios defined as a convex combination of a *sample portfolio* and a scaled *target portfolio*:

$$w_{sh} = (1 - \alpha) w_{sp} + \alpha \nu w_{tg}, \tag{2.13}$$

where w_{sp} is the sample estimator of the true optimal portfolio w_{op}, w_{tg} is the target portfolio, α is the shrinkage intensity, and ν is a scale parameter that we adjust to minimize the bias of the target portfolio.

[5] For a given ϕ, we estimate the portfolio variance using cross-validation as follows. Let us define a sample with T observations. Then, we first delete the i-th observation from our estimation sample. Second, we compute the minimum-variance portfolio from the new sample with $T - 1$ observations. This portfolio is computed with the shrinkage covariance matrix Σ_{sh} calibrated with the method defined in (B.2). Third, we evaluate that portfolio with the i-th observation, which was dropped out of the estimation sample. This is considered the i-th out-of-sample portfolio return. To compute the portfolio variance, we repeat the previous steps with the whole sample, obtaining a time series of T out-of-sample portfolio returns. We estimate the portfolio variance as the sample variance of the out-of-sample portfolio returns.

We consider three shrinkage portfolios obtained by shrinking the sample mean-variance portfolio towards the sample minimum-variance portfolio, the sample mean-variance portfolio towards the equally-weighted portfolio, and the sample minimum-variance portfolio towards the equally-weighted portfolio. Variants of these three shrinkage portfolios have been considered before by Kan and Zhou (2007), Tu and Zhou (2011), and DeMiguel, Garlappi, and Uppal (2009), but there are two main differences between our analysis and the analysis in these papers. First, we introduce an additional scaling parameter that we adjust to minimize the bias of the target portfolio. The advantage of introducing this scaling parameter is that by reducing the bias of the target, we also reduce the overall quadratic loss of the resulting shrinkage portfolio, and our empirical results show that, in general, this improves the out-of-sample performance of the shrinkage portfolios. Second, unlike previous work, we also show how the optimal shrinkage intensity can be estimated using nonparametric techniques.

We consider four different calibration criteria to compute the optimal shrinkage intensity. In addition to the utility maximization criterion and the variance minimization criterion, which are known from the literature, we study two new calibration criteria: the expected quadratic loss minimization criterion, and the Sharpe ratio maximization criterion. Mathematically, we define each method as follows:

$$\text{Expected quadratic loss (eql):} \quad \min_{\alpha} \quad E\left(f_{\text{ql}}\left(w_{sh}\right)\right) = \min_{\alpha} \quad E\left(\|w_{sh} - w_{op}\|_2^2\right), \tag{2.14}$$

$$\text{Utility (ut):} \quad \max_{\alpha} \quad E\left(f_{\text{ut}}\left(w_{sh}\right)\right) = \max_{\alpha} \quad E\left(w_{sh}'\mu - \frac{\gamma}{2} w_{sh}'\Sigma w_{sh}\right), \tag{2.15}$$

$$\text{Variance (var):} \quad \min_{\alpha} \quad E\left(f_{\text{var}}\left(w_{sh}\right)\right) = \min_{\alpha} \quad E\left(w_{sh}'\Sigma w_{sh}\right), \tag{2.16}$$

$$\text{Sharpe ratio (SR):} \quad \max_{\alpha} \quad E\left(f_{\text{SR}}\left(w_{sh}\right)\right) = \max_{\alpha} \quad \frac{E\left(w_{sh}'\mu\right)}{\sqrt{E\left(w_{sh}'\Sigma w_{sh}\right)}}, \tag{2.17}$$

where γ is the investor's risk aversion parameter. The expected utility and Sharpe ratio maximization criteria match the economic incentives of investors and thus the motivation to use them is straightforward.[6] The expected variance minimization criterion also has an economic rationale because investors are often interested in finding those portfolios that minimize the risk of their investments.[7] We consider the expected quadratic loss minimization criterion for two reasons. First, the expected quadratic loss criterion works very well within the context of shrinkage estimators for the covariance matrix; see Ledoit and Wolf (2004b). Thus it is interesting to explore whether it also results in shrinkage portfolios with good performance. Second, the quadratic loss penalizes big errors over small ones, and this in turn is likely to result in more stable portfolio weights with lower

[6]One may consider the Sharpe ratio as the expected value of the ratio between the out-of-sample portfolio mean and the squared root of the out-of-sample portfolio variance; i.e. $E\left(w_{sh}'\mu / \sqrt{w_{sh}'\Sigma w_{sh}}\right)$. However, this expression is not tractable and to approximate it one needs to develop the Taylor expansion of the ratio inside the expectation. Hence, for tractability reasons we define the Sharpe ratio as the ratio between the expected out-of-sample portfolio mean and the squared root of the expected out-of-sample portfolio variance.

[7]In addition, it is knows from the literature that the estimation error in the mean is so large, that it is often more effective to focus on minimizing the variance of portfolio returns; see, for instance, Jagannathan and Ma (2003).

turnover. Nevertheless, in practice the distribution of asset returns may vary with time, which implies that the true optimal portfolio w_{op} may also vary with time, and thus the quadratic loss criterion might fail to provide stable shrinkage portfolios. Presumably, the quadratic loss criterion may be more suitable for shrinkage portfolios that ignore the vector of means, which is more likely to change with time than the covariance matrix.

Characterizing the optimal shrinkage intensity

The following proposition characterizes the optimal shrinkage intensity α for the four calibration criteria. For the expected quadratic loss, utility, and variance criteria, the optimal shrinkage intensity can be obtained in closed-form, whereas for the Sharpe ratio criterion, the optimal shrinkage intensity is the maximizer to an optimization problem and it has to be solved numerically.

Proposition 2. *If asset returns are iid, then the shrinkage intensities for the optimal combination between the sample portfolio and the scaled target portfolio are:*

$$\alpha_{eql} = \frac{E\left(\|w_{sp} - w_{op}\|_2^2\right) - \tau_{sp-tg}}{E\left(\|w_{sp} - w_{op}\|_2^2\right) + E\left(\|\nu w_{tg} - w_{op}\|_2^2\right) - 2\tau_{sp-tg}}, \tag{2.18}$$

$$\alpha_{ut} = \frac{E\left(\sigma_{sp}^2\right) - \nu E\left(\sigma_{sp,tg}\right) - \frac{1}{\gamma}\left(E(\mu_{sp}) - \nu E(\mu_{tg})\right)}{E\left(\sigma_{sp}^2\right) + \nu^2 E\left(\sigma_{tg}^2\right) - 2\nu E\left(\sigma_{sp,tg}\right)}, \tag{2.19}$$

$$\alpha_{var} = \frac{E\left(\sigma_{sp}^2\right) - \nu E\left(\sigma_{sp,tg}\right)}{E\left(\sigma_{sp}^2\right) + \nu^2 E\left(\sigma_{tg}^2\right) - 2\nu E\left(\sigma_{sp,tg}\right)}, \tag{2.20}$$

$$\alpha_{SR} = arg\max_{\alpha} \frac{(1-\alpha)E(\mu_{sp}) + \alpha\nu E(\mu_{tg})}{\sqrt{(1-\alpha)^2 E\left(\sigma_{sp}^2\right) + \alpha^2\nu^2 E\left(\sigma_{tg}^2\right) + 2(1-\alpha)\alpha\nu E\left(\sigma_{sp,tg}\right)}}, \tag{2.21}$$

where $\tau_{sp-tg} = E\left((w_{sp} - w_{op})'(\nu w_{tg} - w_{op})\right)$, $E\left(\sigma_{sp}^2\right) = E\left(w_{sp}'\Sigma w_{sp}\right)$ *is the expected sample portfolio variance,* $E\left(\sigma_{tg}^2\right) = E\left(w_{tg}'\Sigma w_{tg}\right)$ *is the expected target portfolio variance,* $E\left(\sigma_{sp,tg}\right) = E\left(w_{sp}'\Sigma w_{tg}\right)$ *is the expected covariance between the sample portfolio and the target portfolio,* $E\left(\mu_{sp}\right) = E\left(w_{sp}'\mu\right)$ *is the expected sample portfolio mean return, and* $E\left(\mu_{tg}\right) = E\left(w_{tg}'\mu\right)$ *is the expected target portfolio mean return.*

A couple of comments are in order. First, note that, roughly speaking, the optimal shrinkage intensity is the ratio of the error of the sample portfolio, in terms of the specific calibration criterion, divided by the total error of the sample portfolio and the scaled target portfolio.

Second, from (2.19) and (2.20), we observe that the optimal shrinkage intensities of the utility and variance criteria satisfy:

$$\alpha_{ut} = \alpha_{var} - \frac{\frac{1}{\gamma}\left(E(\mu_{sp}) - \nu E(\mu_{tg})\right)}{E\left(\sigma_{sp}^2\right) + \nu^2 E\left(\sigma_{tg}^2\right) - 2\nu E\left(\sigma_{sp,tg}\right)}.$$

This implies that when the expected return of the sample portfolio is larger than the expected return of the scaled target portfolio ($E(\mu_{sp}) > \nu E(\mu_{tg})$), the utility criterion results in a smaller shrinkage intensity than the variance criterion. This is likely to occur when the sample portfolio is the mean-variance portfolio because it is (theoretically) more profitable than the minimum-variance and the equally-weighted portfolios. Under these circumstances, the utility criterion results in more aggressive shrinkage estimators (closer to the sample portfolio) than the variance criterion. This property of the utility criterion may backfire in practice as it is notoriously difficult to estimate mean returns from historical return data. Our empirical results in Section 2.6 confirm this by showing that, when the sample portfolio is the sample mean-variance portfolio, the variance criterion produces better out-of-sample performance than the utility criterion.

2.3. Parametric calibration

In this section, we characterize in closed-form the expectations required to compute the optimal shrinkage intensities under the assumption that returns are iid normal. The closed-form expressions give better insight about the impact of estimation error as a function of the number of assets and observations.

2.3.1. Parametric calibration of the shrinkage moments

We now provide closed-form expressions for the expectations required to calibrate the shrinkage estimators for the covariance and inverse covariance matrices under the assumption that returns are iid normal.

Proposition 3. *Assume that asset returns are iid normal and $T > N + 4$. Moreover, let us define the estimated inverse covariance matrix as $\Sigma_u^{-1} = \frac{T-N-2}{T-1}\Sigma_{sp}^{-1}$, which is the unbiased estimator of the inverse covariance matrix. Hence, the expected quadratic losses of the estimated covariance and inverse covariance matrices are:*

$$E\left(\|\Sigma_{sp} - \Sigma\|_F^2\right) = \frac{N}{T-1}\left(\frac{trace\left(\Sigma^2\right)}{N} + N\left(\overline{\sigma^2}\right)^2\right) \tag{2.22}$$

$$E\left(\left\|\Sigma_u^{-1} - \Sigma^{-1}\right\|_F^2\right) = trace\left(\Omega\right) - trace\left(\Sigma^{-2}\right), \tag{2.23}$$

and

$$E\left(< \Sigma_u^{-1} - \Sigma^{-1}, \nu_{\Sigma^{-1}} I - \Sigma^{-1} >\right) = 0, \tag{2.24}$$

where $\overline{\sigma^2} = trace\left(\Sigma\right)/N$ and $\Omega = \frac{(T-N-2)}{(T-N-1)(T-N-4)}\left(trace\left(\Sigma^{-1}\right)\Sigma^{-1} + (T-N-2)\Sigma^{-2}\right)$.

Note that the expected quadratic loss of the sample estimators increases with the number of assets and decreases with the number of observations. Also, note that we have modified the expression of the estimated inverse covariance matrix to obtain an unbiased estimator. This transformation can only be applied under the normality assumption. However, for the nonparametric approach studied in Section 2.4, we estimate the inverse covariance matrix as the inverse of the sample covariance matrix.

Notice that in order to compute the true optimal shrinkage intensity, we need the population moments of asset returns. In our empirical tests in Section 2.6 we instead use their sample counterparts to estimate the shrinkage intensity, which should also bear some estimation risk. Regardless of the estimation error within the estimated shrinkage intensity, (Tu and Zhou, 2011, Table 5) show that this error is small in the context of shrinkage portfolios, and therefore the estimated optimal shrinkage may outperform the sample portfolio. Our empirical results show that it is also the case for shrinkage moments applied in the context of portfolio optimization.

2.3.2. Parametric calibration of shrinkage portfolios

We now give closed-form expressions for the expectations required to compute the optimal shrinkage intensities given in Proposition 2 for the three shrinkage portfolios and four calibration criteria. In Section 2.5 we exploit these closed-form expressions to improve our understanding of how the impact of estimation error depends on the number of assets N and the number of observations T.

Proposition 4. *Assume returns are independent and normally distributed with mean μ and covariance matrix Σ, and let $T > N+4$. Assume we use the following unbiased estimator of the inverse covariance matrix $\Sigma_u^{-1} = \frac{T-N-2}{T-1}\Sigma_{sp}^{-1}$, and let us construct the sample mean-variance portfolio as $w_{sp}^{mv} = (1/\gamma)\Sigma_u^{-1}\mu_{sp}$, and the sample minimum-variance portfolio as $w_{sp}^{min} = \Sigma_u^{-1}\iota$. Then, the expectations required to compute the optimal shrinkage intensities are given by the following closed-form expressions:*

The expected quadratic loss of the sample mean-variance portfolio:

$$E\left(\left\|w_{sp}^{mv} - w_{op}^{mv}\right\|_2^2\right) = \frac{a}{\gamma^2}\left[trace\left(\Sigma^{-1}\right)\left(\frac{(T-2)}{T} + \mu'\Sigma^{-1}\mu\right) + (T-N-2)\mu'\Sigma^{-2}\mu\right] - \frac{1}{\gamma^2}\mu'\Sigma^{-2}\mu. \tag{2.25}$$

The expected quadratic loss of the sample minimum-variance portfolio with respect to the true mean-variance portfolio:

$$E\left(\left\|\nu w_{sp}^{min} - w_{op}^{mv}\right\|_2^2\right) = \nu^2 a\left[trace\left(\Sigma^{-1}\right)\iota'\Sigma^{-1}\iota + (T-N-2)\iota'\Sigma^{-2}\iota\right] + \frac{1}{\gamma^2}\mu'\Sigma^{-2}\mu - 2\frac{\nu}{\gamma}\iota'\Sigma^{-2}\mu. \tag{2.26}$$

The expected quadratic loss of the sample minimum-variance portfolio:

$$E\left(\left\|w_{sp}^{min} - w_{op}^{min}\right\|_2^2\right) = a\left[trace\left(\Sigma^{-1}\right)\iota'\Sigma^{-1}\iota + (T-N-2)\iota'\Sigma^{-2}\iota\right] - \iota'\Sigma^{-2}\iota. \tag{2.27}$$

The expected value of the sample mean-variance portfolio variance:

$$E\left(\sigma_{mv}^2\right) = E\left(w_{sp}^{mv\prime}\Sigma w_{sp}^{mv}\right) = \frac{1}{\gamma^2}\left(a(T-2)\left(\frac{N}{T} + \mu'\Sigma^{-1}\mu\right)\right). \tag{2.28}$$

The expected value of the sample minimum-variance portfolio variance:

$$E\left(\sigma_{min}^2\right) = E\left(w_{sp}^{min\prime}\Sigma w_{sp}^{min}\right) = a(T-2)\iota'\Sigma^{-1}\iota. \tag{2.29}$$

The expected value of the covariance between the sample mean-variance and sample *minimum-variance portfolios:*

$$E\left(\sigma_{mv,min}\right) = E\left(w_{sp}^{mv\prime}\Sigma w_{sp}^{min}\right) = a(T-2)\frac{1}{\gamma}\mu'\Sigma^{-1}\iota. \quad (2.30)$$

The term τ^2_{mv-min}*:*

$$\begin{aligned}\tau_{mv-min} &= E\left(\left(w_{sp}^{mv} - w_{op}^{mv}\right)'\left(\nu w_{sp}^{min} - w_{op}^{mv}\right)\right)\\ &= \nu\left(\frac{a}{\gamma}\left[trace\left(\Sigma^{-1}\right)\mu'\Sigma^{-1}\iota + (T-N-2)\mu'\Sigma^{-2}\iota\right] - \frac{1}{\gamma}\mu'\Sigma^{-2}\iota\right), \quad (2.31)\end{aligned}$$

where $a = \frac{(T-N-2)}{(T-N-1)(T-N-4)}$*. Moreover, when asset returns are normally distributed, terms* τ_{mv-ew} *and* τ_{min-ew} *are equal to zero,* $E(w_{sp}^{mv})'\mu = (w_{op}^{mv})'\mu$*, and* $E(w_{sp}^{min})'\mu = (w_{op}^{min})'\mu$*.*

Proposition 2 shows that, across every calibration criterion, the shrinkage intensity is higher when the expected quadratic loss or the expected portfolio variance of the sample portfolio are high. Proposition 4 shows that this is likely to occur when the sample covariance matrix is nearly singular. To see this, note that Σ^{-1} and Σ^{-2} appear in the expressions for the quadratic loss and variance of the mean-variance and minimum-variance portfolios, and a nearly singular covariance matrix results in large inverse covariance matrices.

Furthermore, we also see that the expected quadratic loss or the expected portfolio variance of the sample portfolio might be high when we have a low number of observations T compared with the number of assets N. On the other hand, we observe that a small ratio N/T reduces the expected quadratic loss and the portfolio variance. For instance, equation (2.25) converges to zero when the ratio N/T converges to zero. Also, formula (2.28) converges to the true variance of the mean-variance portfolio when the ratio N/T converges to zero.

2.4. Nonparametric calibration of shrinkage estimators

In this section, we describe an alternative nonparametric bootstrap procedure to estimate the optimal shrinkage intensities. We assume that stock returns are iid, but we do not impose any other assumptions on the distribution. Efron (1979) introduces the bootstrap to study the distributional properties of any statistic of interest. Similarly, we use the bootstrap to approximate those expected values of the loss functions required to construct the optimal shrinkage intensities.[8]

This methodology is very intuitive: we generate B bootstrap samples by drawing observations with replacement from the original sample. Then, for each bootstrap sample, we compute the statistic of interest. Finally, we take the sample average among the B bootstrap statistics as an approximation to the expected value.

Contrary to the "simplest" version of bootstrap, we add an error term for each drawn observation. This is what is called *smoothed* bootstrap. We use the multivariate version of the smoothed bootstrap proposed by (Efron, 1979, page 7), such that each extracted

[8]Notice that we do not apply this technique to estimate the shrinkage intensity of the shrinkage vector of means. This is because our proposed technique in Section 2.2.1 is already a nonparametric technique that makes no assumption on the return distribution.

observation is defined as follows:

$$\widetilde{X}_i^* = \mu_{sp} + (I + \Sigma_Z)^{-1/2}\left[X_i^* - \mu_{sp} + \Sigma_{sp}^{1/2} Z_i\right], \tag{2.32}$$

where I is the identity matrix, X_i^* is the i-th randomly drawn observation from $X \in \mathbb{R}^{T\times N}$, μ_{sp} is the sample vector of means of X, Σ_{sp} is the sample covariance matrix of X, and Z_i is a multivariate random variable having zero vector of means and covariance matrix Σ_Z. In the empirical analysis, we set Z_i as a multivariate normal distribution with zero mean and covariance matrix Σ_{sp}, where Σ_{sp} is the sample covariance matrix. The algorithm to compute the optimal shrinkage intensities with the bootstrap analogue is:

Step 1. Construct a bootstrap sample $[X_1^*, X_2^*, \ldots, X_T^*]$ by drawing random observations with replacement from the original sample.

Step 2. Apply formula (2.32) to each drawn observation.

Step 3. Replace the population moments with the original sample moments and compute the corresponding loss function with the bootstrap sample.[9]

Step 4. Repeat Steps 1–3 B times.

Step 5. Average the B bootstrap loss functions to approximate the expected value of the considered loss function.

A positive feature of this technique is that $\widetilde{X}_i^*$ is a random variable which has mean μ_{sp} and covariance matrix Σ_{sp} under the empirical distribution $\widehat{F}$. Another advantage of using the smoothed bootstrap is that we draw observations from a continuous density function, instead of drawing from the set of sample observations, and in turn the probability of having repeated observations is zero. The advantage is that in this manner we avoid the singularity in the estimated covariance matrix, which is likely to occur when there are many repeated observations.

Finally, we have also tested other nonparametric methods like the Jackknife or the d-Jackknife,[10] but we find that the results are not as good as those from using the smoothed bootstrap, and we do not report the results to conserve space.

2.5. Simulation Results

To understand the properties of the shrinkage moments and the shrinkage portfolios, we run a simulation experiment with return data generated by simulating from an iid multivariate normal distribution with sample moments calibrated to those of the 48 industry portfolio dataset from Ken French's website. Under iid normal returns, we can compute

[9] Imagine that we want to approximate $E\left(\|\Sigma_{sp} - \Sigma\|_F^2\right)$. Then, for each bootstrap sample, we compute $\left\|\Sigma_{sp}^b - \Sigma_{sp}\right\|_F^2$ where Σ_{sp}^b is the sample covariance matrix of the b bootstrap sample. Finally, we approximate $E\left(\|\Sigma_{sp} - \Sigma\|_F^2\right) \simeq (1/B)\sum_{b=1}^{B}\left\|\Sigma_{sp}^b - \Sigma_{sp}\right\|_F^2$. Notice that we are replacing Σ_{sp} with Σ_{sp}^b, and Σ with Σ_{sp}.

[10] For a detailed treatment of Jackknife techniques see Efron and Gong (1983) and Efron and Tibshirani (1993). For an application in finance see Basak et al. (2009).

the true optimal shrinkage intensities using the closed-form expressions introduced in Section 2.3. We then use simulated data to characterize the Sharpe ratio of those portfolios computed with shrinkage moments or shrinking the portfolio weights.[11]

Figure 2.1 shows how the true optimal shrinkage intensities for the moments and the Sharpe ratios of the portfolios obtained from these shrinkage moments change with the number of observations. Panel (a) depicts the shrinkage intensities for the Jorion (1986) vector of means and the shrinkage vector of means considered in Section 2.2.1. We observe that the shrinkage intensity of our considered estimator is larger than that of Jorion's estimator. The shrinkage intensity in both estimators represents a measure of the suitability (inadequacy) of the shrinkage target (sample estimator). Although the shrinkage intensities are obtained under different criteria,[12] based on the meaning of shrinkage intensity, we can conclude that the scaled shrinkage target studied in Section 2.2.1 is more suitable than the shrinkage target of Jorion (1986).

Panel (b) in Figure 2.1 gives the shrinkage intensities for the covariance and inverse covariance matrices. This panel shows the optimal shrinkage intensities for both matrices: the shrinkage intensity that minimizes the expected quadratic loss, together with the optimal shrinkage intensity for the covariance matrix that we calibrate by considering both the expected quadratic loss and the condition number. Our first observation is that when we take the condition number into account, we obtain a larger shrinkage intensity than when we focus solely on the expected quadratic loss. Our second observation is that the shrinkage intensity for the inverse covariance matrix is quite large, specially in small samples. This is because for small samples, the sample covariance matrix is nearly singular, and thus the impact of estimation error explodes when we invert the sample covariance matrix. Consequently, the expected quadratic loss criterion results in very large shrinkage intensities.

Panel (c) in Figure 2.1 depicts the Sharpe ratios for the mean-variance portfolios formed with the Jorion (1986) vector of means, and our considered shrinkage estimator of means. For the data simulated from a multivariate normal distribution, the Sharpe ratio of the mean-variance portfolio with the Jorion (1986) vector of means is larger than the Sharpe ratio of the mean-variance portfolio formed with our considered shrinkage estimator of means. The reason for this is that our proposed estimator results in a larger shrinkage intensity, and hence the scaled target has a greater importance in the resulting estimated mean. In turn, the obtained mean-variance portfolio exploits less the difference of expected returns across assets, which gives a less profitable portfolio. However, this method may help to provide more stable mean-variance portfolios when the sample vector of means is difficult to estimate. Indeed, our results in Section 2.6 show that this conservative approach works well when applied to the empirical data due to the instability of the sample vector of means.

Panel (d) in Figure 2.1 depicts the simulated Sharpe ratios for the minimum-variance portfolios formed with the shrinkage estimators of the covariance matrix and the inverse

[11] We simulate 5,000 samples of length T, and for each sample we compute all the considered shrinkage estimators using the true optimal shrinkage intensities. We compute the desired portfolio and we compute the out-of-sample portfolio return and the out-of-sample portfolio variance of each portfolio. We approximate the expected portfolio return and the expected portfolio variance with the sample average among the 5,000 generated values. We use the estimated expected portfolio return and the estimated expected portfolio variance to compute the Sharpe ratios.

[12] The shrinkage intensity of the Jorion (1986) vector of means is obtained from an empirical-Bayes approach, whereas our studied estimator chooses the shrinkage intensity to minimize the expected quadratic loss.

covariance matrix. We observe that the minimum-variance portfolio formed with the shrinkage covariance matrix that accounts for both the expected quadratic loss and the condition number attains the largest Sharpe ratio. This suggests that even for a sample size of $T = 250$ observations, it is important to take into account the singularity of the sample covariance matrix, and thus the condition number matters to calibrate the shrinkage intensity. On the other hand, we observe that the minimum-variance portfolio formed with the shrinkage inverse covariance matrix attains the lowest Sharpe ratio. As mentioned before, minimizing the expected quadratic loss of the shrinkage inverse covariance matrix gives a very large shrinkage intensity that results in a portfolio too close to the equally-weighted portfolio.

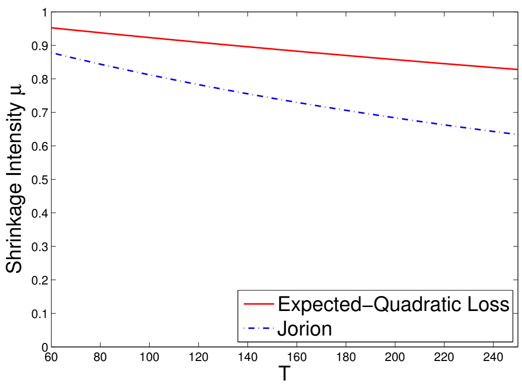
Figure 2.2 depicts the shrinkage intensities and Sharpe ratios for the shrinkage portfolios calibrated with the methods described in Section 2.2.2. Panels (a) and (b) give the shrinkage intensities for the mv-min and mv-ew shrinkage portfolios, obtained by shrinking the sample mean-variance portfolio towards the minimum-variance and equally-weighted portfolios, respectively. For both shrinkage portfolios, the variance minimization criterion provides the largest shrinkage intensity. The reason for this is that, for both portfolios, the shrinkage targets are low-variance portfolios. Because the variance minimization criterion is not a utility-maximizing criterion, the portfolios calibrated with this criterion attain the lowest Sharpe ratios, as shown in panels (d) and (e).

Panel (c) depicts the shrinkage intensities for the min-ew shrinkage portfolio, obtained by shrinking the minimum-variance portfolio towards the equally-weighted portfolio. Since the expected quadratic loss minimization criterion seeks the stability of portfolio weights (see Section 2.2.2), and the equally weighted portfolio is a rather stable portfolio, this calibration criterion provides the largest shrinkage intensity. The stability of portfolio weights does not guarantee low portfolio variance and/or high expected return, as this is more dependent on market conditions. Therefore, this calibration criterion provides a slightly lower Sharpe ratio than the other calibration criteria, as we observe in panel (f).

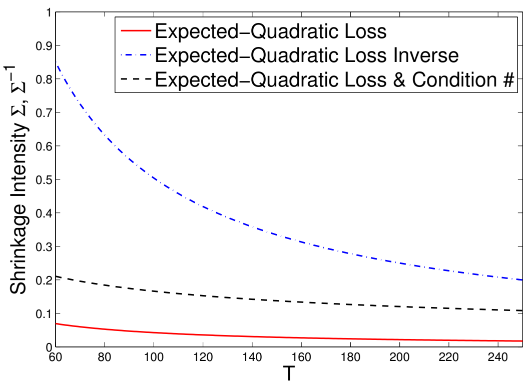
Comparing panels (d), (e), and (f) in Figure 2.2, we see that among the shrinkage portfolios, the mv-min shrinkage portfolio obtains the worst Sharpe ratio for small samples. This is because this shrinkage portfolio is obtained from the sample mean-variance and sample minimum-variance portfolios, which contain substantial estimation error for small samples. Moreover, we see that there always exists a combination which beats the sample portfolio in terms of Sharpe ratio, although for large samples the difference becomes smaller.

Figure 2.1.: Shrinkage intensities and Sharpe ratios of portfolios computed with shrinkage moments

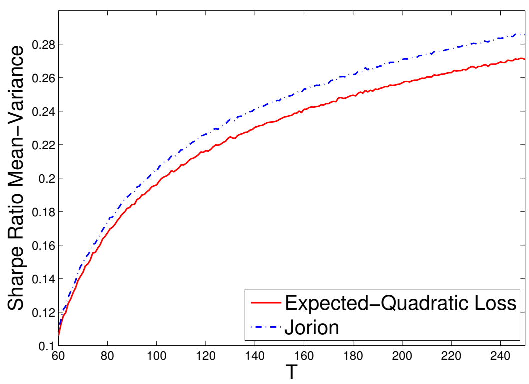
These plots show the evolution of the true optimal shrinkage parameters for the shrinkage estimators of μ, Σ and Σ^{-1}, as well as the Sharpe ratios of portfolios formed with the shrinkage moments. Plot (a) depicts the evolution of the shrinkage intensities for the vector of means of our studied shrinkage mean vector of returns (solid line) and the Jorion (1986) mean vector of returns (dot-dashed line). Plot (b) depicts the shrinkage intensities of the shrinkage covariance studied in Section 2.2.1 (solid line), the shrinkage inverse covariance matrix studied in Section 2.2.1 (dot-dashed line), and the shrinkage covariance matrix studied in Section 2.2.1 (dashed line). Plot (c) depicts the simulated Sharpe ratios of the mean-variance portfolios constructed with our studied shrinkage mean vector of returns (solid line), and the Jorion (1986) mean vector of returns (dot-dashed line). Plot (d) depicts the simulated Sharpe ratios of the minimum-variance portfolios constructed with the shrinkage covariance matrix studied in Section 2.2.1 (solid line), the shrinkage inverse covariance matrix studied in Section 2.2.1 (dot-dashed line), and the shrinkage covariance matrix studied in Section 2.2.1 (dashed line). To carry out the simulation, we use the sample moments of a dataset formed by 48 industry portfolios (48IndP) as the population moments of a multivariate normal distribution. The shrinkage estimator for the covariance matrix accounting for its expected quadratic loss and its condition number establishes $\phi = 100$. The experiment is made considering an investor with a risk aversion parameter of $\gamma = 10$.



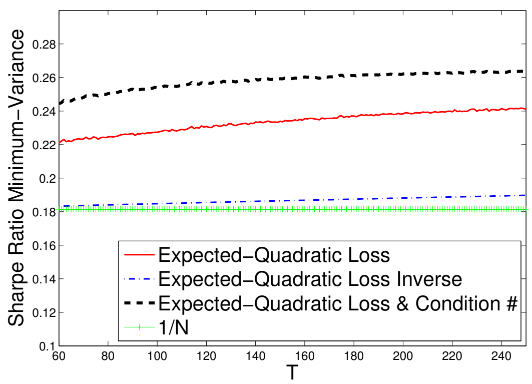
(a) Shrinkage intensity of μ_{sh}



(b) Shrinkage intensity of Σ_{sh} and Σ_{sh}^{-1}



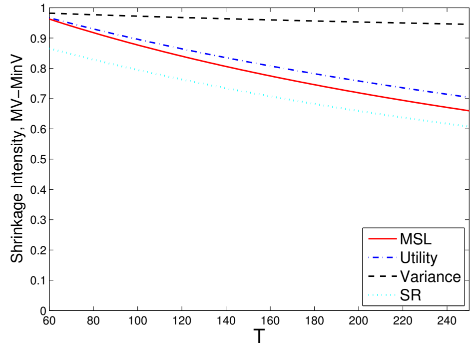
(c) Sharpe ratio of w_{sp}^{mv} formed with μ_{sh}



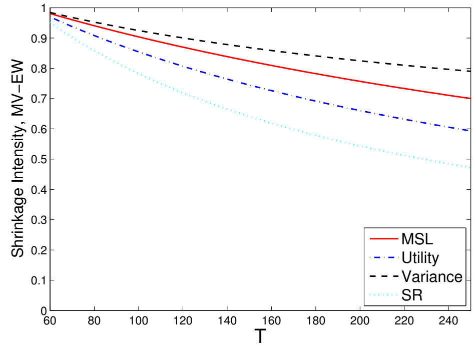
(d) Sharpe ratio of w_{sp}^{min} formed with Σ_{sh} and Σ_{sh}^{-1}

Figure 2.2.: Shrinkage intensities and Sharpe ratios of shrinkage portfolios

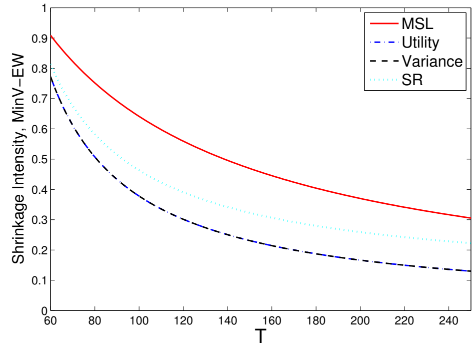
These plots show the evolution of the true optimal shrinkage parameters for the shrinkage portfolios, as well as their Sharpe ratios. For each shrinkage portfolio, we compute the corresponding considered value (the shrinkage intensity or the Sharpe ratio) under every calibration criterion, where EQL, Utility, Variance and SR stand for the expected quadratic loss minimization criterion (solid line), utility maximization criterion (dot-dashed line), variance minimization criterion (dashed line), and Sharpe ratio maximization criterion (dotted line), respectively. On the other hand, Mv and Min, in plots (d)-(f), stand for the Sharpe ratios of the sample mean-variance portfolio and the sample minimum-variance portfolio, respectively. To carry out the simulation, we use the sample moments of a dataset formed by 48 industry portfolios (48IndP) as the population moments of a multivariate normal distribution. The experiment is made considering an investor with a risk aversion parameter of $\gamma = 10$.



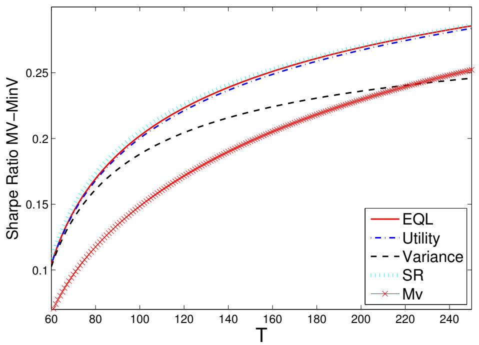
(a) Shrinkage parameter evolution for mv-min



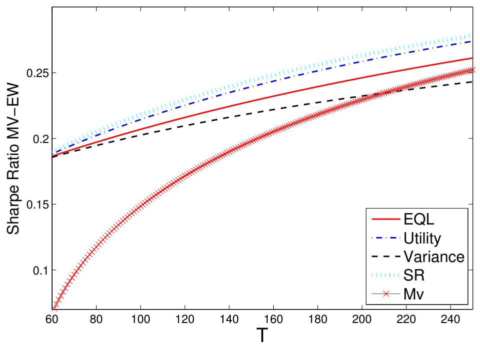
(b) Shrinkage parameter evolution for mv-ew



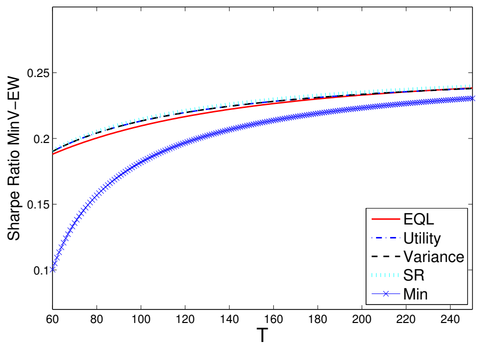
(c) Shrinkage parameter evolution for min-ew



(d) Sharpe ratio of mv-min



(e) Sharpe ratio of mv-ew



(f) Sharpe ratio of min-ew

We also study the case where the number of observations is lower than the number of assets.[13] This case is relevant in practice for portfolio managers who deal with a large number of assets when the number of return observations that are relevant to the prevailing market conditions is small. We assume that returns follow an iid multivariate normal distribution defined with the sample moments of the 48IndP dataset. We again characterize the Sharpe ratios for the minimum-variance portfolios computed with the shrinkage covariance matrix calibrated by the expected quadratic loss minimization criterion, and also the calibration criterion that takes into account both the expected quadratic loss and the matrix condition number.

Figure 2.3.: Sharpe ratios of portfolios formed with shrinkage covariance matrices
This plot shows the evolution of the Sharpe ratios for the minimum-variance portfolios composed with the shrinkage covariance matrix studied in Section 2.2.1 (dashed line with rhombus) and the shrinkage covariance matrix studied in Section 2.2.1 (dot-dashed line). For the sake of comparison, we also plot the results of the equally-weighted portfolio (solid line). To compute the shrinkage covariance matrix studied in Section 2.2.1, we use $\phi = 100$, as in the previous simulations. To carry out the simulation, we use the sample moments of a dataset formed by 48 industry portfolios (48IndP) as the population moments of a multivariate normal distribution.

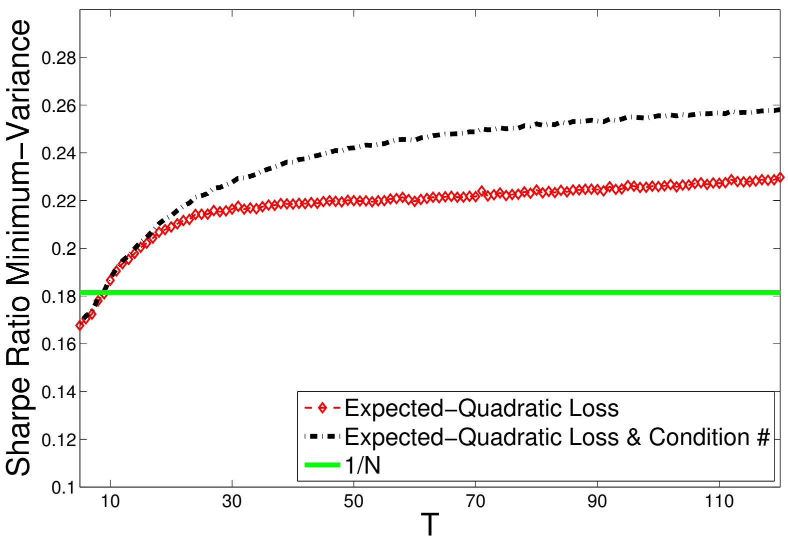


Figure 2.3 depicts the results of the experiment. When the sample size is very small, e.g. $T = 10$, the expected quadratic loss of the covariance matrix is very large. Consequently, the shrinkage intensity that minimizes the expected quadratic loss is high. Therefore, the resulting shrinkage estimator of the covariance matrix has a reasonable condition number. This is why, for very small samples, both shrinkage methods provide similar shrinkage intensities and, in turn, similar Sharpe ratios.[14] On the other hand, when the sample size is bigger than 20 observations, the Sharpe ratios of the minimum-variance portfolios computed with the shrinkage estimators from Sections 2.2.1 and 2.2.1 diverge. This is because for larger sample sizes ($T > 20$), the expected quadratic loss of the sample covariance matrix is lower and therefore, the shrinkage intensity that minimizes

[13]For this part of the analysis, we only study minimum-variance portfolios computed from shrinkage covariance matrices, which are not singular.

[14]Notice that these results depend on parameter ϕ, which establishes the trade-off between expected quadratic loss and condition number.

the quadratic loss is relatively small. As a result, although this shrinkage intensity is sufficient to reduce the quadratic loss, it is not large enough to keep the condition number small. Hence, by taking the condition number explicitly into consideration we therefore can improve the performance of the resulting portfolios for large samples.

In general, we observe that the minimum-variance portfolios formed with the shrinkage covariance matrix studied in Section 2.2.1 with $\phi = 100$ have larger Sharpe ratios than the minimum-variance portfolio formed with the shrinkage covariance matrix studied in Section 2.2.1, specially for sample sizes larger than 20 observations. Therefore, we conclude that it is always beneficial to account for the matrix condition number, which is specially useful for managers dealing with large number of assets.

Finally, to study the robustness of our results with respect to the number of assets, we repeat our simulations for the case where the number of observations is fixed to $T = 150$, but the number of assets changes. The robustness check analysis is made across five different datasets (5IndP, 10IndP, 38IndP, 48IndP and 100FF) listed in Table 2.1. We observe from the results, which we do not report to conserve space, that the insights from our experiment are robust to the number of assets.

2.6. Empirical Results

Table 2.1 lists the six datasets considered in the analysis. We consider 4 industry portfolio datasets from Ken French's website. These are portfolios of all stocks from NYSE, AMEX and NASDAQ grouped in terms of their industry. We use datasets with stocks grouped into 5, 10, 38, and 48 industries (5IndP, 10IndP, 38IndP, 48IndP). We also consider a dataset of 100 portfolios formed from stocks sorted by size and book-to-market ratio (100FF), downloaded from Ken French's website. The last dataset (SP100) is formed by 100 stocks, randomly chosen the first month of each new year from the set of assets in the S&P500 for which we have returns for the entire estimation window, as well as for the next twelve months.

Table 2.2 lists all the portfolios considered. Panel A lists the portfolios from the existing literature that we consider as benchmarks. The first benchmark portfolio is the classical mean-variance portfolio of Markowitz (1952).[15] The second portfolio is the classical mean-variance portfolio composed with the shrinkage vector of means proposed by Jorion (1986). The next three portfolios are mixtures of portfolios proposed in the literature; the first one is the mixture of the mean-variance and minimum-variance portfolio of Kan and Zhou (2007); the second is the mixture of the mean-variance and equally-weighted portfolios studied by Tu and Zhou (2011); the third is the mixture of the minimum-variance and equally-weighted portfolio of DeMiguel, Garlappi, and Uppal (2009). The sixth portfolio is the minimum-variance portfolio. The seventh portfolio is the minimum-variance portfolio formed with the shrinkage covariance matrix of Ledoit and Wolf (2004b), which shrinks the sample covariance matrix to the identity matrix. The eighth portfolio is the minimum-variance portfolio formed with the shrinkage covariance matrix of Ledoit and Wolf (2003), which shrinks the sample covariance matrix to the sample covariance matrix of a single-index factor model. The ninth portfolio is the equally-weighted portfolio. Panel B lists the portfolios constructed with the shrinkage estimators studied in Section 2.2.1. The first portfolio in Panel B is the mean-variance portfolio with the shrinkage vector of means studied in Section 2.2.1. The second portfolio is the minimum-variance

[15]For our empirical evaluation, we set the risk aversion coefficient $\gamma = 5$

Table 2.1.: List of Datasets

This table list the various datasets analyzed, the abbreviation used to identify each dataset, the number of assets N contained in each dataset, the time period spanned by the dataset, and the source of the data. The dataset of CRSP returns (SP100) is constructed in a way similar to Jagannathan and Ma (2003), with monthly rebalancing: in January of each year we randomly select 100 assets as our asset universe for the next 12 months.

[a] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

[b] CRSP, The Center for Research in Security Prices

#	Dataset	Abbreviation	N	Time Period	Source
1	5 Industry Portfolios representing the US stock market	5Ind	5	01/1972-06/2009	K. French[a]
2	10 Industry Portfolios representing the US stock market	10Ind	10	01/1972-06/2009	K. French
3	38 Industry Portfolios representing the U.S stock market	38IndP	38	01/1972-06/2009	K. French
4	48 Industry Portfolios representing the U.S. stock market	48Ind	48	01/1972-06/2009	K. French
5	100 Fama and French Portfolios of firms sorted by size and book to market	100FF	100	01/1972-06/2009	K. French
6	100 randomized stocks from S&P 500	SP100	100	01/1988-12/2008	CRSP[b]

portfolio formed with the shrinkage covariance matrix studied in Section 2.2.1. The third portfolio is the minimum-variance portfolio formed with the shrinkage inverse covariance matrix proposed in Section 2.2.1. The fourth portfolio is the minimum-variance portfolio formed with a shrinkage covariance matrix calibrated by accounting for the expected quadratic loss and the condition number. The shrinkage covariance matrices of the last three portfolios are calibrated under the parametric approach, assuming normality, and under the bootstrap nonparametric approach. Panel C lists the shrinkage portfolios proposed in Section 2.2.2. In the empirical analysis, we calculate the shrinkage intensities of these portfolios using the four calibration methods defined in Section 2.2.2. Again, we compute the shrinkage intensities under a parametric approach, and also under a bootstrap nonparametric approach.[16]

2.6.1. Out-of-sample performance evaluation

We compare the out-of-sample performance of the different portfolios with two different criteria: (i) out-of-sample portfolio Sharpe ratio adjusted with transaction costs, and (ii) out-of-sample portfolio standard deviation.[17] We use the "rolling-horizon" procedure to

[16]For the nonparametric approach, we generate B=500 bootstrap samples. We have also run the empirical application with B=1000 and B=2000 bootstrap samples, but the results are similar to the case of B=500 samples.

[17]We also computed the turnover but because the Sharpe ratio is adjusted with transaction costs, we do not report the results for the turnover to conserve space. These results are provided in a Supplementary Appendix.

Table 2.2.: List of portfolio models

This table lists the various portfolio strategies considered in the empirical study. Panel A lists the existing portfolios from the literature. Panel B lists portfolios where the moments are shrunk with the methods proposed in Section 2.2. Panel C lists the shrinkage portfolio. The third column gives the abbreviation that we use to refer to each strategy.

#	Policy	Abbreviation
	Panel A: Benchmark portfolios	
1	Classical mean-variance portfolio	mv
2	Bayes-Stein mean-variance portfolio	bs
3	Kan-Zhou's (2007) three-fund portfolio	kz
4	Mixture of mean-variance and equally-weighted (Tu and Zhou (2011))	tz
5	Mixture of minimum-variance and equally-weighted DeMiguel et.al. (2009))	dm
6	Minimum-Variance portfolio	min
7	Minimum-variance portfolio with Ledoit and Wolf (2004) shrinkage covariance matrix, which shrinks the sample covariance matrix to the identity matrix	lw
8	Minimum-variance portfolio with Ledoit and Wolf (2003) shrinkage covariance matrix, which shrinks the sample covariance matrix to the sample covariance matrix of a single-index factor model	lw-m
9	Equally-weighted portfolio	1/N or ew
	Panel B: Portfolios estimated with new calibration procedures to shrink moments	
	Shrinkage mean-variance portfolio	
10	Mean-variance portfolio formed with the shrinkage vector of means defined in Section 2.2.1	f-mv
	Shrinkage minimum-variance portfolio	
11	Formed with Ledoit and Wolf (2004) shrinkage covariance matrix: calibrated under a parametric calibration assuming normality and calibrated under a bootstrap nonparametric approach	par-lw and npar-lw
12	Formed with the shrinkage inverse covariance matrix studied in Section 2.2.1: calibrated under a parametric calibration assuming normality and calibrated under a bootstrap nonparametric approach	par-ilw and npar-ilw
13	Formed with a shrinkage covariance matrix that accounts for the expected quadratic loss and the condition number: calibrated under a parametric calibration assuming normality and calibrated under a bootstrap nonparametric approach	par-clw and npar-clw
	Panel C: Shrinkage portfolios	
14	Mixture of mean-variance and scaled minimum-variance portfolios	mv-min
15	Mixture of mean-variance and scaled equally-weighted portfolios	mv-ew
16	Mixture of minimum-variance and scaled equally-weighted portfolios	min-ew

compute the out-of-sample performance measures. The "rolling-horizon" is defined as follows: first, we choose a window over which to estimate the portfolio. The length of the window is $M < T$, where T is the total number of observations of the dataset. In the empirical analysis, our estimation window has a length of $M = 120$, which corresponds with

10 years of data (with monthly frequency). Second, we compute the various portfolios using the return data over the estimation window. Third, we repeat the "rolling-window" procedure for the next month by including the next data point and dropping the first data point of the estimation window. We continue doing this until the end of the dataset. Therefore, at the end we have a time series of $T-M$ portfolio weight vectors for each of the portfolios considered in the analysis; that is $w_t^i \in \mathbb{R}^N$ for $t = M, \dots, T-1$ and portfolio i.

The out-of-sample returns are computed by holding the portfolio weights for one month w_t^i and evaluate it with the next-month vector of excess returns: $r_{t+1}^i = R_{t+1}' w_t^i$, where R_{t+1} denotes the vector of excess returns at time $t+1$ and r_{t+1}^i is the out-of-sample portfolio return at time $t+1$ of portfolio i. We use the times series of portfolio returns and portfolio weights of each strategy to compute the out-of-sample standard deviation and Sharpe ratio:

$$(\sigma^i)^2 = \frac{1}{T-M-1} \sum_{t=M}^{T-1} \left(w_t^{i'} R_{t+1} - \overline{r}^i \right)^2, \tag{2.33}$$

$$\text{with } \overline{r}^i = \frac{1}{T-M} \sum_{t=M}^{T-1} \left(w_t^{i'} R_{t+1} \right), \tag{2.34}$$

$$SR^i = \frac{\overline{r}^i}{\sigma^i}, \tag{2.35}$$

where w_t^i is the vector of weights at t under policy i. To account for transaction costs in the empirical analysis, the definition of portfolio return is slightly corrected by the implied cost of rebalancing the portfolio. Then, the definition of portfolio return, net of proportional transaction costs, is:

$$\underline{r}_{t+1}^i = (1 + R_{t+1}' w_t^i) \left(1 - \kappa \sum_{j=1}^{N} \left| w_{j,t+1}^i - w_{j,t^+}^i \right| \right) - 1, \tag{2.36}$$

where $w_{j,t}^i$ denotes the estimated portfolio weight of asset j at time t under policy i, w_{j,t^+}^i is the estimated portfolio weight of asset j accumulated at time $t+1$, and κ is the chargeable fee for rebalancing the portfolio. In the empirical analysis, expressions (3.24)-(3.23) are computed using portfolio returns discounted by transaction costs.

Finally, to measure the statistical significance of the difference between the adjusted Sharpe ratios, we use the stationary bootstrap of Politis and Romano (1994) with B=1000 bootstrap samples and block size b=1.[18] We use the methodology suggested in (Ledoit and Wolf, 2008, Remark 2.1) to compute the resulting bootstrap p-values. Furthermore, we also measure the statistical significance of the difference between portfolio variances by computing the bootstrap p-values using the methodology proposed in Ledoit and Wolf (2011).

2.6.2. Discussion of the out-of-sample performance

Table B.3 reports the annualized Sharpe ratio adjusted by transaction costs of the benchmark portfolios and the portfolios constructed with the shrinkage estimators studied in

[18] We have also computed the p-values when b=5. The interpretation of the results does not change for b=1 or b=5.

Section 2.2.1. We consider transaction costs of 50 basis points—Balduzzi and Lynch (1999) argue that 50 basis points is a good estimate of transaction costs for an investor who trades with individual stocks. Panel A of Table B.3 reports the Sharpe ratios for the benchmark portfolios. We observe that the minimum-variance portfolio with the shrinkage covariance matrix proposed by Ledoit and Wolf (2004b) (lw) attains the highest out-of-sample Sharpe ratio among all benchmark portfolios. Panel B reports the Sharpe ratio for the portfolios formed with the shrinkage estimators studied in Section 2.2.1 calibrated under the assumption of iid normal returns. We observe that the minimum-variance portfolio formed from the shrinkage covariance matrix that accounts for the expected quadratic loss and the condition number (par-clw) outperforms the lw portfolio for medium and large datasets. This is because for medium and large datasets, the sample covariance matrix is more likely to be nearly singular, and in turn it is important to control for the condition number to construct optimal portfolios. Furthermore, we can observe that the differences between par-clw and lw are statistically significant for the 38IndP, 48IndP and 100FF datasets. Consequently, for medium and large datasets it is significantly relevant to use a calibration criterion that explicitly takes into consideration the condition number of the covariance matrix.

Panel C reports the portfolios constructed with the shrinkage moments calibrated with the proposed smoothed bootstrap of Section 2.4.[19] First, we observe that the mean-variance portfolio obtained from the shrinkage vector of means studied in Section 2.2.1 beats the benchmark mean-variance portfolios (mv and bs) for small and medium datasets; i.e. $N \leq 48$ assets. We also observe that, in general, the smoothed bootstrap approach works better than the parametric approach to calibrate the shrinkage covariance matrix of minimum-variance portfolios. Hence, the proposed bootstrap approach works effectively with empirical datasets where the available observations depart from the normality assumption.

Table B.4 reports the annualized Sharpe ratio adjusted by transaction costs of the shrinkage portfolios studied in Section 2.2.2. Panel A reports the annualized adjusted Sharpe ratio of the shrinkage portfolios calibrated via parametric assumptions; see Section 2.3. Panel B reports the annualized adjusted Sharpe ratio of the shrinkage portfolios calibrated via bootstrap; see Section 2.4. Panel C reports the results of the shrinkage portfolios from the literature. From Panel A, we make two observations. First, the variance minimization criterion is the best criterion in small and medium datasets for portfolios that consider the vector of means, mv-min and mv-ew, whereas the expected quadratic loss is the best calibration criterion for the portfolio that does not consider the vector of means, min-ew. This result confirms the intuition about this criterion discussed in Section 2.2.2.

From Panel B in Table B.4 we observe that, in general, the best shrinkage portfolio is the mixture formed with the minimum-variance portfolio and the equally-weighted portfolio. We also observe that the expected quadratic loss minimization criterion is, in general, the best calibration criterion in terms of Sharpe ratio and the results obtained under the nonparametric bootstrap approach are slightly better than those obtained under the assumption of normally distributed returns. This is because empirical returns depart from the normality assumption and our proposed smoothed bootstrap approach captures this characteristic.

Panel C of Table B.4 shows the annualized Sharpe ratio of the existing mixture of port-

[19]For the vector of means, we use the criterion proposed in Section 2.2.1 because it does not require any parametric assumption.

Table 2.3.: Annualized Sharpe ratio of benchmark portfolios and portfolios with shrinkage moments (κ =50 basis points)

This table reports the out-of-sample annualized Sharpe ratio of benchmark portfolios and portfolios constructed by using the shrinkage estimators studied in Section 2.2. We adjust the Sharpe ratio with transaction costs, where we assume that transaction costs are equal to 50 basis points (bp). We consider an investor with a risk aversion parameter of $\gamma = 5$. One, two and three asterisks indicate that the difference with the lw portfolio is statistically different from zero for a 90%, 95% and 99% confidence interval, respectively.

Policy	5IndP	10IndP	38IndP	48IndP	100FF	SP100
Panel A: Benchmark Portfolios						
Portfolios that consider the vector of means						
mv	0.591**	0.519***	-0.023***	-0.154***	0.187***	-0.386***
bs	0.807	0.765***	0.203***	0.024***	0.169***	-0.330***
Portfolios that do not consider the vector of means						
min	0.895	0.934	0.478***	0.343***	-0.959***	-0.188***
lw	0.893	0.961	0.752	0.662	0.954	0.602
lw-m	0.890	0.953	0.670***	0.623	0.804***	0.589
Naïve Portfolios						
1/N	0.786	0.817	0.717	0.712	0.754	0.340
Panel B: Portfolios calibrated parametrically						
Portfolios that do not consider the vector of means						
par-lw	0.891	0.944	0.645***	0.559***	0.687***	0.563
par-ilw	0.902	0.918	0.730	0.715	0.755	0.341
par-clw	0.890	0.956	0.823**	0.792***	1.194***	0.622
Panel C: Portfolios calibrated nonparametrically						
Portfolios that consider the vector of means						
f-mv	0.854	0.797**	0.374***	0.255***	-0.960***	-0.241***
Portfolios that do not consider the vector of means						
npar-lw	0.888**	0.957*	0.729***	0.634***	0.870***	0.586
npar-ilw	0.864	0.867	0.721	0.713	0.754	0.340
npar-clw	0.888	0.965	0.859***	0.831***	1.178***	0.622

folios from the literature. We observe that among these portfolios, the mixture formed by the minimum-variance portfolio and the equally weighted portfolio offers the best results for small and medium datasets; i.e. $N \leq 48$ assets. This mixture, however, performs worse than our studied shrinkage portfolio formed with the minimum-variance portfolio and the equally-weighted portfolio across every dataset. Hence, in general our proposed framework to construct shrinkage portfolios turns out to hedge better the investor's portfolio against estimation error.

Tables B.5 and B.6 report the results for the out-of-sample standard deviation of the studied portfolios. These results are consistent with the results for the Sharpe ratio adjusted by transaction costs. First, we observe that our proposed shrinkage vector of means provides mean-variance portfolios with lower variance than those computed with the Bayes-Stein estimator of Jorion (1986) across every dataset. Second, we observe that the condition number helps to obtain minimum-variance portfolios with lower variability than minimum-variance portfolios computed with the shrinkage covariance matrix of

Table 2.4.: Annualized Sharpe ratio with transaction costs of shrinkage portfolios (κ =50 basis points)

This table reports the out-of-sample annualized Sharpe ratio (adjusted with 50 bp) of the studied shrinkage portfolios for an investor with $\gamma = 5$. One, two and three asterisks indicate that the difference with the lw portfolio is statistically different from zero for a 90%, 95% and 99% confidence interval, respectively.

Policy	5IndP	10IndP	38IndP	48IndP	100FF	SP100
Panel A: Shrinkage portfolio with parametric calibration						
EQL Minimization						
mmv-min	0.780	0.755***	0.285***	0.180***	-1.333***	-0.222***
mv-ew	0.713**	0.778*	0.439**	0.316**	0.195***	-0.073**
min-ew	0.887	0.949	0.642	0.590	-0.285	0.007
Utility Maximization						
mv-min	0.764	0.726	0.314	0.269	-0.989	-0.206
mv-ew	0.732	0.732	0.413	0.394	0.205	0.031
min-ew	0.888	0.944	0.564	0.460	-0.719***	-0.118***
Variance Minimization						
mv-min	0.892	0.921	0.479***	0.330***	-1.071***	-0.213***
mv-ew	0.764**	0.806	0.617	0.680	-0.332***	0.072*
min-ew	0.888	0.944	0.563***	0.460***	-0.718***	-0.118***
Sharpe Ratio Maximization						
mv-min	0.784	0.756**	0.248***	0.093***	0.187***	-0.230**
mv-ew	0.698*	0.665**	0.263***	0.145***	-1.213***	-0.119***
min-ew	0.843	0.948	0.567***	0.486***	-0.718***	-0.055***
Panel B: Shrinkage portfolios with bootstrap calibration						
EQL Minimization						
mv-min	0.747*	0.744***	-0.315***	-0.208***	-0.503***	0.307
mv-ew	0.723*	0.794*	0.582	0.616	0.689	0.340
min-ew	0.885	0.949	0.715	0.704	0.754	0.340
Utility Maximization						
mv-min	0.735*	0.723***	-0.233***	-0.653***	-0.377***	-0.208**
mv-ew	0.747*	0.761**	0.582	0.668	0.737	0.340
min-ew	0.858*	0.940	0.704	0.709	0.754	0.340
Variance Minimization						
mv-min	0.872	0.925	0.450***	0.339***	-0.975***	-0.195***
mv-ew	0.767*	0.801	0.681	0.707	0.754	0.340
min-ew	0.849**	0.940	0.703*	0.714	0.754	0.340
Sharpe Ratio Maximization						
mv-min	0.739	0.739***	0.178***	-0.174***	0.184***	-0.357***
mv-ew	0.696*	0.677**	0.372**	0.452	0.755	0.340
min-ew	0.844	0.947	0.601	0.632	0.755	0.341
Panel C: Existing mixture of portfolios						
kz	0.784	0.756***	0.248***	0.093***	0.184***	-0.230***
tz	0.708*	0.693***	0.341***	0.358*	-0.569***	0.082**
dm	0.825	0.905	0.564***	0.486***	-0.733***	-0.055***

Ledoit and Wolf (2004b), specially for medium and large datasets where the estimated covariance matrix is more likely to be near singular. For the shrinkage portfolios that consider the vector of means, mv-min and mv-ew, we observe that the variance criterion tends to provide portfolios with lower variability, whereas the shrinkage portfolio that does not consider the vector of means has a lower variability when it is calibrated with the expected quadratic loss criterion. Comparing the parametric and nonparametric approaches, we observe that generally, the smoothed bootstrap approach provides as good results as the parametric approach, and in many cases the smoothed bootstrap approach gives portfolios with lower variability.

Table 2.5.: Standard deviation of benchmark portfolios and portfolios with shrinkage moments

This table reports the out-of-sample standard deviation of benchmark portfolios and portfolios constructed by using the shrinkage estimators studied in Section 2.2. We consider an investor with a risk aversion parameter of $\gamma = 5$. One, two and three asterisks indicate that the difference with the lw portfolio is statistically different from zero for a 90%, 95% and 99% confidence interval, respectively.

Policy	5IndP	10IndP	38IndP	48IndP	100FF	SP100
Panel A: Benchmark Portfolios						
Portfolios that consider the vector of means						
mv	0.175**	0.167***	0.277***	0.414***	4.777***	0.578***
bs	0.146	0.137**	0.185***	0.273***	3.930***	0.495***
Portfolios that do not consider the vector of means						
min	0.139	0.127	0.134***	0.150***	0.260***	0.307***
lw	0.137	0.124	0.123	0.129	0.128	0.129
lw-m	0.138	0.126	0.124***	0.127	0.133***	0.128
Naïve Portfolios						
1/N	0.153	0.148	0.166	0.165	0.173	0.170
Panel B: Portfolios calibrated parametrically						
Portfolios that do not consider the vector of means						
par-lw	0.137	0.125	0.126***	0.134***	0.142***	0.132
par-ilw	0.139	0.133	0.160	0.162	0.172	0.169
par-clw	0.137	0.124	0.122**	0.125***	0.121***	0.126
Panel C: Portfolios calibrated nonparametrically						
Portfolios that consider the vector of means						
f-mv	0.140	0.134**	0.147***	0.193***	0.322***	0.345***
Portfolios that do not consider the vector of means						
npar-lw	0.137**	0.124*	0.124***	0.131***	0.132***	0.130
npar-ilw	0.144	0.140	0.164	0.164	0.173	0.170
npar-clw	0.137	0.123	0.121**	0.124***	0.121***	0.126

We now summarize the main findings from our empirical analysis. Our first observation is that portfolios computed from the shrinkage vector of means calibrated by minimizing its expected quadratic loss outperform those computed from the Bayes-Stein vector of means of Jorion (1986). Second, we observe that controlling for the condition number of the shrinkage covariance matrix results in portfolio weights that are more stable, and this leads to better adjusted Sharpe ratios for medium and large datasets. Third, for

Table 2.6.: Standard deviation of shrinkage portfolios

This table reports the out-of-sample standard deviation of the studied shrinkage portfolios. We consider an investor with a risk aversion parameter of $\gamma = 5$. One, two and three asterisks indicate that the difference with the lw portfolio is statistically different from zero for a 90%, 95% and 99% confidence interval, respectively.

Policy	5IndP	10IndP	38IndP	48IndP	100FF	SP100
Panel A: Shrinkage portfolios with parametric calibration						
EQL Minimization						
mv-min	0.149	0.138***	0.164***	0.198***	0.383***	0.348**
mv-ew	0.155*	0.140*	0.171*	0.187**	1.866***	0.245**
min-ew	0.138	0.125	0.129	0.139	0.173***	0.185**
Utility Maximization						
mv-min	0.151	0.139***	0.158***	0.187**	0.275***	0.312***
mv-ew	0.153	0.142**	0.174**	0.191*	1.475***	0.207**
min-ew	0.138	0.126	0.129***	0.142***	0.211***	0.236**
Variance Minimization						
mv-min	0.139	0.128	0.134***	0.151***	0.270***	0.310***
mv-ew	0.149*	0.140	0.162	0.164	0.204***	0.198**
min-ew	0.138	0.126	0.129***	0.142***	0.211***	0.236**
Sharpe Ratio Maximization						
mv-min	0.149	0.138**	0.174***	0.236***	2.469***	0.362**
mv-ew	0.156*	0.148**	0.195***	0.252***	0.446***	0.276**
min-ew	0.139	0.127	0.130***	0.140***	0.223***	0.226**
Panel B: Shrinkage portfolios with nonparametric calibration						
EQL Minimization						
mv-min	0.150*	0.138***	1.323***	0.903***	0.601***	29.758
mv-ew	0.155*	0.140	0.161	0.159	0.175	0.170
min-ew	0.138	0.126	0.137	0.152	0.173	0.170
Utility Maximization						
mv-min	0.152*	0.138***	0.526***	0.726***	0.423***	0.321**
mv-ew	0.152	0.140**	0.161	0.160	0.173	0.170
min-ew	0.139*	0.126	0.133	0.150	0.173	0.170
Variance Minimization						
mv-min	0.139	0.128	0.135***	0.151***	0.261***	0.308**
mv-ew	0.150*	0.142	0.162	0.163	0.173	0.170
min-ew	0.140**	0.126	0.134	0.150	0.173	0.170
Sharpe Ratio Maximization						
mv-min	0.151*	0.139***	0.201***	0.298***	4.860***	0.501***
mv-ew	0.156*	0.147**	0.177**	0.177	0.172	0.170
min-ew	0.140	0.127	0.132*	0.144	0.172	0.170
Panel C: Existing mixture of portfolios						
kz	0.149	0.138***	0.174***	0.235***	2.243***	0.362***
tz	0.155**	0.145**	0.184**	0.208*	0.242***	0.194*
dm	0.139	0.129	0.130***	0.140***	0.223***	0.226**

shrinkage portfolios that consider the vector of means, the variance minimization criterion is the most robust criterion, whereas for shrinkage portfolios that do not consider the vector of means, the expected quadratic loss criterion works better. Finally, the studied nonparametric approach to calibrate shrinkage estimators captures the departure from normality in real return data and this results in more stable portfolios (small turnover) with reasonable Sharpe ratios.

2.7. Summary

We provide a comprehensive investigation of shrinkage estimators for portfolio selection. We first study several shrinkage estimators of the moments of asset returns. We propose a new calibration criterion for the shrinkage estimator of the vector of means and we obtain a closed-form expression of the true optimal shrinkage intensity without making any assumptions on the distribution of stock returns. This new calibration criterion for the shrinkage vector of means turns out to perform better than the vector of means proposed by Jorion (1986). We also introduce a novel criterion to calibrate the shrinkage covariance matrix proposed by Ledoit and Wolf (2004b). This new calibration criterion accounts for both the expected quadratic loss and the condition number of the covariance matrix. Our empirical results show that the shrinkage estimator based on this criterion results in portfolios with larger Sharpe ratio, adjusted with transaction costs, and lower standard deviation for medium and large datasets.

For shrinkage portfolios, we consider two novel calibration criteria (expected quadratic loss and Sharpe ratio) in addition to the expected utility criterion, considered in most of the existent literature, and the variance minimization criterion considered in DeMiguel, Garlappi, and Uppal (2009). Our empirical results show that the variance minimization criterion is the most robust to calibrate shrinkage portfolios that make use of the sample vector of means. On the other hand, the expected quadratic loss minimization criterion is the most robust procedure to calibrate those portfolios that ignore the vector of means.

Finally, we show that the smoothed bootstrap approach is a practical and simple technique to calibrate shrinkage estimators in situations where the available data departs from the normality assumption. In general, we observe that portfolios computed using this approach perform well in medium and large datasets.

To the best of our knowledge, this work is among the first to consider and compare different shrinkage estimators within the context of portfolio optimization. This chapter attempts to highlight the importance of calibrating shrinkage estimators to construct optimal portfolios. In the empirical application we demonstrate that the results may be very different when using alternative calibration criteria, and we show that the size of the shrinkage intensity matters in the out-of-sample performance of optimal portfolios constructed from shrinkage estimators.

3. Parameter Uncertainty in Multiperiod Portfolio Optimization with Transaction Costs

3.1. Overview

The seminal paper of Markowitz (1952) shows that an investor who cares only about the portfolio mean and variance should hold one of the portfolios on the efficient frontier. Markowitz's mean-variance framework is the main foundation of most practical investment approaches, but it relies on three restrictive assumptions. First, the investor is *myopic* and maximizes a one-period utility. Second, financial market are frictionless. Third, the investor knows the exact parameters that capture asset price dynamics. In this chapter, we study the case where these three assumptions fail to hold; that is, the investor tries to maximize a *multiperiod* utility in the presence of *quadratic transaction costs* and suffers from parameter uncertainty. Our contribution is threefold. First, we characterize analytically the utility loss associated with estimation error for a multiperiod mean-variance investor who faces quadratic transaction costs. Second, we use these result to propose two shrinkage portfolios designed to combat the impact of parameter uncertainty. Third, we provide evidence based on simulated and empirical datasets that the proposed shrinkage portfolios substantially outperform the portfolios of investors that ignore either parameter uncertainty or transaction costs.

There is an extensive literature on multiperiod portfolio selection in the presence of transaction costs under the assumption that there is *no parameter uncertainty.* For the case with a single-risky asset and proportional transaction costs, Constantinides (1979) and Davis and Norman (1990) show that the optimal portfolio policy of an investor with constant relative risk aversion (CRRA) utility is characterized by a no-trade region. The case with multiple-risky assets and proportional transaction costs is generally intractable analytically.[1] Garleanu and Pedersen (2012) show that the case with multiple-risky assets and *quadratic* transaction costs is, however, more tractable; and they provide closed-form expressions for the optimal portfolio policy of a multiperiod mean-variance investor.[2]

As we observe in Chapter 2, there is also an extensive literature on parameter uncertainty on portfolio selection for the case of a myopic investor who *is not subject to transaction costs*. Kan and Zhou (2007) characterize analytically the utility loss of a mean-variance investor who suffers from parameter uncertainty. Moreover, they consider a three-fund portfolio, which is a combination of the sample mean-variance portfolio, the sample minimum-variance portfolio, and the risk-free asset. They analytically characterize those combination weights of three-fund portfolios that minimize the investor's utility

[1] Liu (2004), however, characterizes analytically the case where asset returns are uncorrelated for the particular case of an investor with constant absolute risk aversion (CARA) utility.

[2] Quadratic transaction costs are well suited to model market impact cost; see, for instance, Engle and Ferstenberg (2007).

loss from parameter uncertainty.[3]

Our work is, to the best of our knowledge, the first to consider the impact of parameter uncertainty on the performance of a multiperiod mean-variance investor facing quadratic transaction costs. As mentioned above, our contribution is threefold. Our first contribution is to give a closed-form expression for the utility loss of an investor who uses sample information to construct her optimal portfolio policy. We find that the utility loss is the product of two terms. The first term is the single-period utility loss in the absence of transaction costs, as characterized by Kan and Zhou (2007). The second term captures the effect of the multiperiod horizon on the overall utility loss. Specifically, this term can be split into the losses from the multiperiod mean-variance utility and the multiperiod transaction costs.

We also use our characterization of the utility loss to understand how the transaction costs and the investor's impatience factor affect the investor utility loss. We observe that agents that face high transaction costs are less affected by estimation risk. Although high trading costs do not diminish the investor's exposure to estimation risk, they delay its impact to future stages where the overall importance in the investor's expected utility is lower. Also, an investor with high impatience factor is less affected by estimation risk. Roughly speaking, the investor's impatience factor has a similar effect on the investor's expected utility to that of trading costs. When the investor is more impatient, the cost of making a trade takes a greater importance than the future expected payoff of the corresponding trade. Hence, larger trading costs or higher impatience factor make the investor trade less aggressively, and this offsets the uncertainty of the inputs that define the multiperiod portfolio model.

Our second contribution is to propose shrinkage portfolios designed to combat estimation risk in the multiperiod mean-variance framework with quadratic transaction costs. From Garleanu and Pedersen (2012), it is easy to show that, in the absence of estimation error, the optimal portfolio policy is to trade towards the Markowitz portfolio at a fixed trading rate every period. For this reason, we propose two approaches to combat estimation error: i) shrink the Markowitz portfolio maintaining the trading rate fixed at its nominal value; ii) shrink the trading rate. Regarding the first approach i), we propose a shrinkage portfolio that is obtained by shrinking the Markowitz portfolio towards zero. We term this portfolio as multiperiod three-fund portfolio, because it is a combination of the current portfolio, the Markowitz portfolio, and the risk-free asset. Then, we propose a second shrinkage portfolio obtained by shrinking the Markowitz portfolio towards a target portfolio that is less affected by estimation error, and we term the resulting shrinkage portfolio as four-fund portfolio. We show that the shrinkage intensities for the three- and four-fund portfolios are the same as for the single-period investor and we show that it is always optimal to shrink the Markowitz portfolio and combine it with the minimum-variance portfolio. Regarding the second approach ii), the nominal trading rate given by Garleanu and Pedersen (2012) may not be optimal in the presence of parameter uncertainty. Hence, we propose versions of previous four-fund portfolio where the trading rate is also shrunk to reduce the effects of parameter uncertainty. We provide a rule to compute the optimal trading rate and we illustrate those conditions where the investor can obtain gains by shrinking the trading rate.

Our third contribution is to evaluate the out-of-sample performance of the proposed shrinkage portfolios on simulated data as well as on an empirical dataset of commodity

[3]See also Tu and Zhou (2011), who consider a combination of the sample mean-variance portfolio with the equally-weighted portfolio.

futures similar to that used by Garleanu and Pedersen (2012). We find that the four-fund portfolios (either with fixed or optimal trading rate) substantially outperform portfolios that either ignore transaction costs, or ignore parameter uncertainty. In addition, we find that shrinking the nominal trading rate can also improve the investor's out-of-sample performance.

The outline of the chapter is as follows. In Section 3.2, we introduce the setup of the economy, and we characterize the investor's expected loss when the investor uses sample information to construct the trading strategy in Section 3.3. In Section 3.4, we introduce the shrinkage portfolios that help to reduce the effects of estimation risk, and we test their out-of-sample performance in Section 3.5. We summarize our main findings in Section 3.6.

3.2. General framework

We adopt the framework proposed by Garleanu and Pedersen (2012), henceforth the G&P model. In this framework, the investor maximizes her multiperiod mean-variance utility, net of quadratic transaction costs, by choosing the number of shares to hold from each of the N risky assets. The only difference between our model and the G&P model is that while G&P assume that price changes in excess of the risk-free rate are predictable, we focus on the case where price changes are independent and identically distributed (iid) as normal with mean μ and covariance matrix Σ, which is a common assumption in most of the transaction costs literature; see Constantinides (1979), Davis and Norman (1990), Liu and Loewenstein (2002), and Liu (2004).

The investor's objective is

$$\max_{\{x_i\}} \quad U\big(\{x_i\}\big) = \sum_{i=0}^{\infty} (1-\rho)^{i+1} \left(x_i'\mu - \frac{\gamma}{2} x_i'\Sigma x_i \right) - (1-\rho)^i \left(\frac{\lambda}{2} \Delta x_i' \Sigma \Delta x_i \right), \qquad (3.1)$$

where $x_i \in R^N$ for $i \geq 0$ contains the number of shares held from each of the N risky assets at time i, ρ is the investor's impatience factor, and γ is the risk-aversion parameter. The term $(\lambda/2)\Delta x_i'\Sigma\Delta x_i$ is the quadratic transaction cost at the ith period, where λ is the transaction cost parameter, and $\Delta x_i = x_i - x_{i-1}$ is the vector containing the number of shares traded at the ith period.

A few comments are in order. First, quadratic transaction costs are appropriate to model market impact costs, which arise when the investor makes large trades that distort market prices. A common assumption in the literature is that market *price* impact is linear on the amount traded (see Kyle (1985)), and thus market impact costs are quadratic.[4] Second, we adopt G&P's assumption that the quadratic transaction costs are proportional to the covariance matrix Σ. G&P provide micro-foundations to justify this type of trading cost.[5]

[4]Several authors have shown that the quadratic form matches the market impact costs observed in empirical data; see, for instance, Lillo et al. (2003) and Engle et al. (2012).

[5]In addition, Greenwood (2005) shows from an inventory perspective that price changes are proportional to the covariance of price changes. Engle and Ferstenberg (2007) show that under some assumptions, the cost of executing a portfolio is proportional to the covariance of price changes. Transaction costs proportional to risk can also be understood from the dealer's point of view. Generally, the dealer takes at time i the opposite position of the investor's trade and "lays it off" at time $i+1$. In this sense, the dealer has to be compensated for the risk of holding the investor's trade.

It is easy to adapt the results in G&P to obtain a closed-form expression for the optimal portfolio policy in our setting.

Proposition 5 (Adapted from Garleanu and Pedersen (2012))**.** *The optimal portfolio at time i is:*

$$x_i = (1-\beta)\, x_{i-1} + \beta x^M, \tag{3.2}$$

where $x^M = \frac{1}{\gamma}\Sigma^{-1}\mu$ is the static mean-variance (Markowitz) portfolio, $\beta = \frac{\sqrt{(\gamma+\widetilde{\lambda}\rho)^2+4\gamma\lambda}-(\gamma+\widetilde{\lambda}\rho)}{2\lambda}$, $\widetilde{\lambda} = (1-\rho)^{-1}\lambda$, and $\beta \leq 1$ is the trading rate. Moreover, the monotonicity properties of the trading rate β are as follows:

1. *β is monotonically increasing with γ.*
2. *β is monotonically decreasing with λ.*
3. *β is monotonically decreasing with ρ.*

Proposition 5 shows that the optimal portfolio policy is to trade every period at a trading rate β towards the static mean-variance (Markowitz) portfolio. The intuition is that the Markowitz portfolio is optimal in terms of the multiperiod mean-variance utility, but it is prohibitive to trade in a single period to the Markowitz portfolio due to the impact of transaction costs.

3.3. Multiperiod utility loss

In this section, we study the impact of parameter uncertainty by characterizing analytically the investor's expected loss. We consider an investor who uses a *plug-in* approach to estimate the optimal portfolio policy given by Proposition 5. Specifically, let r_l for $l = 1, 2, \ldots, T$ be the sample of excess price changes with which the investor constructs the following unbiased estimator of the Markowitz portfolio: $\widehat{x}^M = \widehat{\Sigma}^{-1}\widehat{\mu}/\gamma$, where

$$\widehat{\mu} = \frac{1}{T}\sum_{l=1}^{T} r_l, \quad \text{and} \quad \widehat{\Sigma} = \frac{1}{T-N-2}\sum_{l=1}^{T}\left(r_l - \widehat{\mu}\right)^2. \tag{3.3}$$

Then, the estimated optimal portfolio policy is given by replacing x^M in (3.2) with $\widehat{x}^M$,

$$\widehat{x}_i = (1-\beta)\widehat{x}_{i-1} + \beta\widehat{x}^M, \tag{3.4}$$

which results in an unbiased estimator of the optimal trading strategy.

Like Kan and Zhou (2007) we define the investor's expected utility loss as the difference between the investor's utility evaluated for the true optimal portfolio and the investor's expected utility evaluated for the estimated portfolio. For a single-period mean-variance investor in the absence of transaction costs, Kan and Zhou (2007) characterize the expected utility loss corresponding to the sample mean-variance portfolio $\hat{x}^M$, which is defined as $\delta_S(x^M, \widehat{x}^M) = U_S(x^M) - E\left[U_S(\widehat{x}^M)\right]$, where $U_S(x^M) = x^{M\prime}\mu - \frac{\gamma}{2}x^{M\prime}\Sigma x^M$:

$$\delta_S(x^M, \widehat{x}^M) = (c-1)\frac{\theta}{2\gamma} + \frac{1}{2\gamma}c\frac{N}{T}, \tag{3.5}$$

where $c = [(T-N-2)(T-2)]/[(T-N-1)(T-N-4)]$.[6] We observe that the expected loss for a static investor decreases with γ and the sample length T, whereas it increases with $\theta = \mu'\Sigma^{-1}\mu$ and the number of available assets N.

The following proposition provides a closed-form expression for the utility loss of a multiperiod mean-variance investor facing quadratic transaction costs that uses the plug-in approach described above.

Proposition 6. *A multiperiod mean-variance investor who uses the plug-in approach to estimate the optimal portfolio policy has the following expected utility loss:*

$$\delta(\{x_i\}, \{\widehat{x}_i\}) = \delta_S(x^M, \widehat{x}^M) \times \underbrace{[AV + AC]}_{\textit{Multiperiod term}}, \tag{3.6}$$

where AV is the multiperiod mean-variance loss factor, and AC is the multiperiod transaction cost loss factor:

$$AV = \frac{1-\rho}{\rho} + \frac{(1-\rho)(1-\beta)^2}{1-(1-\rho)(1-\beta)^2} - 2\frac{(1-\rho)(1-\beta)}{1-(1-\rho)(1-\beta)}, \tag{3.7}$$

$$AC = \frac{\lambda}{\gamma}\frac{\beta^2}{1-(1-\rho)(1-\beta)^2}. \tag{3.8}$$

Proposition 6 shows that the multiperiod utility loss is equal to the single-period utility loss multiplied by the summation of two terms. The first term captures the losses from the multiperiod mean-variance utility, and the second term captures the losses from the multiperiod transaction costs. Note also that the multiperiod loss factors AV and AC depend only on λ, γ, and ρ.

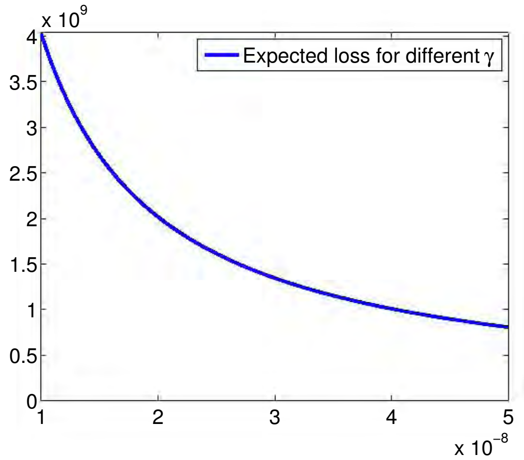
Figure 3.1 depicts the absolute multiperiod expected losses for different values of γ, λ, and ρ. We consider a base-case investor with $\gamma = 10^{-8}$, $\lambda = 3 \times 10^{-7}$ and $\rho = 1 - \exp(-0.1/260)$, which are the same parameters that define our base-case investor in the empirical application in Section 3.5.[7] In addition, the investor constructs the optimal trading strategy with $T = 500$ observations, and we define the population parameters μ and Σ with the sample moments of the empirical dataset of commodity futures used in the empirical application in Section 3.5. We obtain three main findings from Figure 3.1. First, the multiperiod expected loss decreases with γ. Like in the static case, this is a natural result because as the investor becomes more risk averse, the investor's exposure to risky assets is lower, and then the impact of parameter uncertainty is also smaller. Second, the multiperiod expected loss decreases with λ. As trading costs increase, the investor delays the convergence to the Markowitz portfolio and in turn, the investor postpones the impact of parameter uncertainty to future stages where the overall importance of utility losses is smaller. This makes that the multiperiod expected loss becomes smaller with trading costs. Third, the multiperiod expected loss decreases with ρ. Roughly speaking, the investor's impatience factor has a similar effect on the investor's expected utility to that of trading costs. When the investor is more impatient, the cost of making a trade takes a greater importance than the future expected payoff of the corresponding trade.

[6] Expression (3.5) is not the exact expected loss that we find in Kan and Zhou (2007). This has been adapted to our estimator for the covariance matrix, that provides an unbiased estimator of the Markowitz portfolio, whereas the estimate for this element in Kan and Zhou (2007) provides a biased estimator of the Markowitz portfolio.

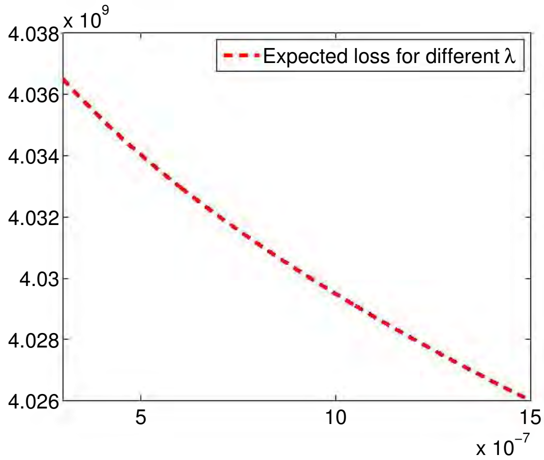
[7] See Section 3.5 to understand the implications from these parameters.

Figure 3.1.: Absolute loss of multiperiod investor

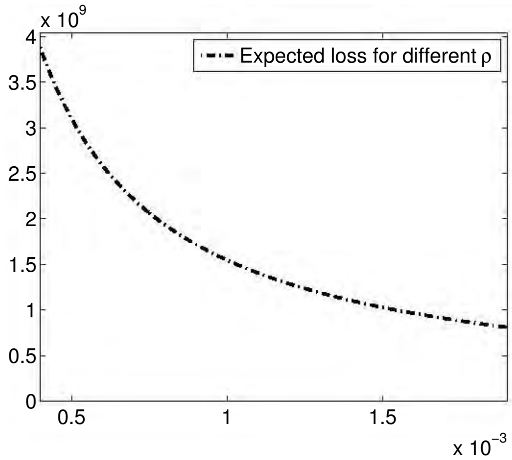
This plot depicts the investor's absolute expected loss for different values of γ, λ, and ρ. Our base-case investor is defined with $\gamma = 10^{-8}$, $\lambda = 3 \times 10^{-7}$ and $\rho = 1 - \exp(-0.1/260)$. We consider an investor that has 500 observations to construct the optimal trading strategy whose parameters are defined with the sample moments of the empirical dataset formed with commodities that we consider in the empirical application.



(a) Different values of γ



(b) Different values of λ

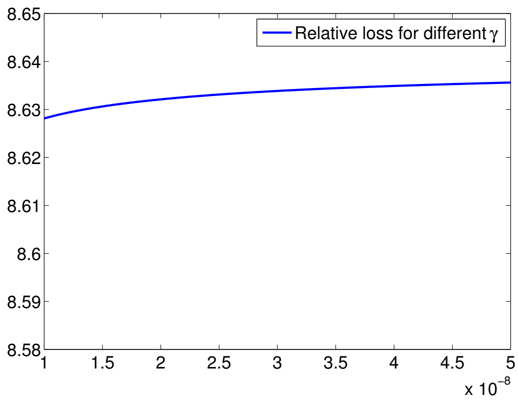


(c) Different values of ρ

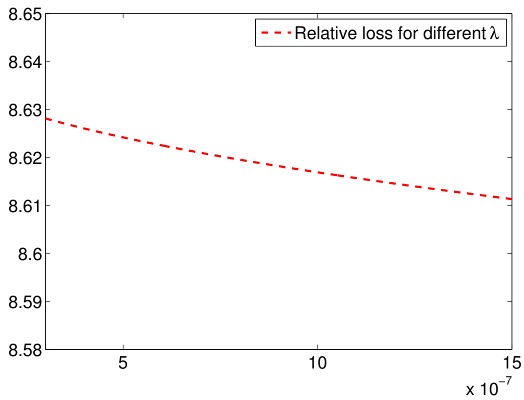
Although the above example gives some monotonicity properties of the absolute utility loss, for interpretation it may be useful to study how the *relative* utility loss depends on the investor's risk aversion parameter γ, trading costs λ, and the investor's impatience factor ρ. Figure 3.2 depicts the investor's relative loss for different values of γ, λ and ρ. From Figure 3.3a, we observe that as the investor's risk aversion parameter increases, the investor's relative loss also increases but slightly. That is, the relative loss is nearly constant (but increasing) with the investor's risk aversion parameter. On the other hand, Figure 3.3b illustrates that larger trading costs reduce the investor's relative loss. Finally, we observe in Figure 3.3c that an investor with high impatience factor has a lower relative loss. In turn, the variation of the investor's utility loss is, in absolute value, lower than that of the investor's utility when the risk aversion parameter changes, whereas the variation of the investor's utility loss is larger than that of the investor's utility when the impatience factor or trading costs change.

Figure 3.2.: Relative loss of multiperiod investor

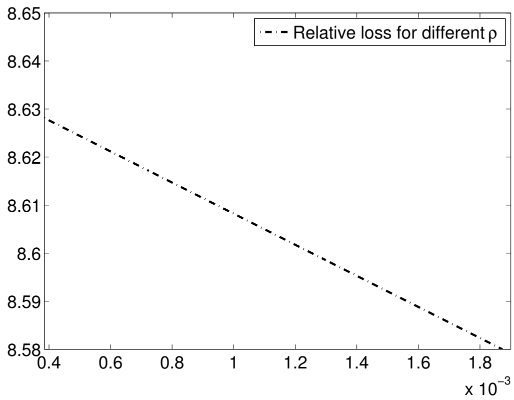
This plot depicts the investor's relative loss for different values of γ, λ, and ρ. Our base-case investor is defined with $\gamma = 10^{-8}$, $\lambda = 3 \times 10^{-7}$ and $\rho = 1 - \exp(-0.1/260)$. We consider an investor that has 500 observations to construct the optimal trading strategy whose parameters are defined with the sample moments of the empirical dataset of commodity futures that we consider in the empirical application.



(a) Different values of γ



(b) Different values of λ



(c) Different values of ρ

After analyzing the expected utility loss of an investor who uses sample information to construct her optimal portfolio, in section 3.4 we propose several shrinkage portfolios that help to reduce the effects of estimation risk on the performance of multiperiod portfolios.

3.4. Multiperiod shrinkage portfolios

In this section we propose several shrinkage portfolios that mitigate the impact of estimation error on the multiperiod mean-variance utility of an investor who faces quadratic transaction costs. We consider two approaches to shrink the plug-in portfolio policy defined in Equation (3.4): (i) shrink the estimated Markowitz portfolio x^M, and (ii) shrink the trading rate β.

3.4.1. Shrinking the Markowitz portfolio

The optimal portfolio at period i, in the absence of estimation error, allocates the investor's wealth into three funds: the risk-free asset, the portfolio at period $i-1$, and the Markowitz portfolio. However, this solution is not optimal when the investor suffers from parameter uncertainty. For the single period case, Kan and Zhou (2007) show that shrinking the Markowitz portfolio helps to mitigate the impact of parameter uncertainty.

We generalize their analysis to the multiperiod case. In particular, we consider two different approaches to shrink the Markowitz portfolio. First, we consider shrinking the Markowitz portfolio towards the portfolio that invests solely on the risk-free asset; that is, towards $x = 0$. We term the resulting shrinkage portfolio as multiperiod three-fund portfolio because the optimal portfolio at period i allocates the investor's wealth into three different funds: the portfolio at time $i-1$, the Markowitz portfolio, and the risk-free asset. The resulting portfolio can be written as:

$$\widehat{x}_i^{3F} = (1-\beta)\widehat{x}_{i-1}^{3F} + \beta\eta\widehat{x}^M, \tag{3.9}$$

where η is the shrinkage intensity.

Second, we consider a multiperiod portfolio that combines the Markowitz portfolio with a target portfolio. This combination may diversify the effects of estimation error in the sample mean-variance portfolio and reduce the risk of taking inefficient positions. We choose as a target portfolio the minimum-variance portfolio $\widehat{x}^{Min} = (1/\gamma)\Sigma^{-1}\iota$, which is known to be less sensitive to estimation error than the mean-variance portfolio.[8] We term the resulting shrinkage portfolio as four-fund portfolio:

$$\widehat{x}_i^{4F} = (1-\beta)\widehat{x}_{i-1}^{4F} + \beta(\varsigma_1\widehat{x}^M + \varsigma_2\widehat{x}^{Min}), \tag{3.10}$$

where ς_1 and ς_2 are the combination parameters for the Markowitz portfolio and the minimum-variance portfolio, respectively.

Note that while Kan and Zhou (2007) consider a static mean-variance investor that is not subject to transaction costs, we consider a multiperiod mean-variance investor subject to quadratic transaction costs. Given this, one would expect that the optimal shrinkage intensities for our proposed multiperiod shrinkage portfolios would differ from those obtained by Kan and Zhou (2007) for the single-period case, but the following proposition shows that the optimal shrinkage intensities for the single-period and multiperiod cases coincide.

Proposition 7. *The optimal shrinkage intensities for the three-fund and four-fund portfolios that minimize the utility loss of a multiperiod mean-variance investor* $\delta(\{x_i\}, \{\widehat{x}_i\})$ *coincide with the optimal shrinkage intensities for the single-period investor who ignores transaction costs. Specifically, the optimal shrinkage intensity for the three-fund portfolio*

[8] Notice that the minimum-variance portfolio does not consider γ. However, for expository reasons, we multiply the unscaled minimum-variance portfolio with $(1/\gamma)$ to simplify the analysis.

η and the optimal combination parameters for the four-fund portfolio ς_1 and ς_2 are:

$$\eta = c^{-1}, \tag{3.11}$$

$$\varsigma_1 = c^{-1}\frac{\Psi^2}{\Psi^2 + \frac{N}{T}}, \tag{3.12}$$

$$\varsigma_2 = c^{-1}\frac{\frac{N}{T}}{\Psi^2 + \frac{N}{T}} \times \frac{\mu'\Sigma^{-1}\iota}{\iota'\Sigma^{-1}\iota}, \tag{3.13}$$

where $c = [(T-2)(T-N-2)]/[(T-N-1)(T-N-4)]$ and $\Psi^2 = \mu'\Sigma^{-1}\iota - (\mu'\Sigma^{-1}\iota)^2/(\iota'\Sigma^{-1}\iota) > 0$.

Note that the optimal shrinkage intensities for the multiperiod three-fund an four-fund portfolios do not depend on transaction costs, given by parameter λ, and as a result they coincide with the optimal shrinkage intensities for the single-period case in the absence of transaction costs.

The following corollary shows that the optimal multiperiod portfolio policy that ignores estimation error is inadmissible in the sense that it is always optimal to shrink the Markowitz portfolio. Moreover, the three-fund shrinkage portfolio is also inadmissible in the sense that it is always optimal to shrink the Markowitz portfolio towards the target minimum-variance portfolio. The result demonstrates that the shrinkage approach is bound to improve performance under our main assumptions.

Corollary 1. *It is always optimal to shrink the Markowitz portfolio; that is, $\eta < 1$. Moreover, it is always optimal to combine the Markowitz portfolio with the target minimum-variance portfolio; that is, $\varsigma_2 > 0$.*

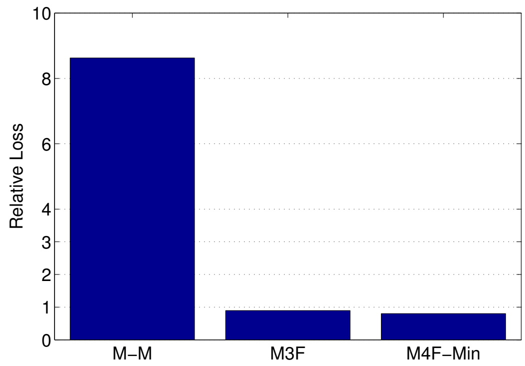
As expected from Corollary 1, the *relative improvement* in the investor's expected utility when using the proposed shrinkage portfolios in (3.9) and (3.10) is larger than that when using the plug-in portfolio in (3.4). In particular, Figure 3.3 shows that for the base-case investor that we consider in Section 3.3, the *relative loss* when using the shrinkage three-fund portfolio in (3.9) is about *eight times smaller* than that when using the plug-in multiperiod portfolio in (3.4). And the *relative loss* when using the shrinkage four-fund portfolio in (3.10) is about *11% less* than that when using the three-fund portfolio in (3.9). Figure 3.3 shows that there is a clear advantage of using the four-fund portfolio with respect to the plug-in multiperiod portfolio and the multiperiod shrinkage three-fund portfolio.

3.4.2. Shrinking the trading rate

In this section we study the additional utility gain associated with shrinking the trading rate in addition to the target portfolio. For the proposed shrinkage portfolios in (3.9) and (3.10), note that the *nominal* trading rate β as given in Proposition 5 may not be optimal in the presence of parameter uncertainty. To mitigate even more this effect, we propose to optimize the trading rate in order to minimize the investor's utility loss from estimation risk. In particular, a multiperiod mean-variance investor who uses the shrinkage four-fund portfolio in (3.10) may reduce the impact of parameter uncertainty by minimizing the corresponding expected utility loss, $\delta(\{x_i\}, \{\widehat{x}_i^{4F}(\beta)\})$, respect to the trading rate β.

Figure 3.3.: Relative loss of different multiperiod investor

This plot depicts the investor's relative loss of the plug-in multiperiod investor (M-M), the multiperiod investor that shrinks the static mean-variance portfolios (M3F), and the multiperiod four-fund portfolio that combines the static mean-variance portfolio with the minimum-variance portfolio (M4F-Min). Our base-case investor is defined with $\gamma = 10^{-8}$, $\lambda = 3 \times 10^{-7}$ and $\rho = 1 - \exp(-0.1/260)$. The investor has 500 observations to construct the optimal trading strategy whose parameters are defined with the sample moments of the empirical dataset of commodity futures that we consider in the empirical application.



The following proposition formulates an equivalent optimization problem to obtain the optimal trading rate for the shrinkage four-fund portfolio in (3.10). Notice that we can apply the same proposition to the shrinkage three-fund portfolio in (3.9) simply by considering $\varsigma_2 = 0$ and $\varsigma_1 = \eta$.

Proposition 8. *For the shrinkage four-fund portfolio in* (3.10)*, the optimal trading rate β that minimizes the expected utility loss $\delta(\{x_i\}, \{\widehat{x}_i^{4F}(\beta)\})$ can be obtained by solving the following optimization problem:*

$$\max_{\beta} \quad \overbrace{V_1(x_{-1} - x^C)'\mu}^{\text{Excess return}} - \frac{1}{2}\overbrace{\left(E\left[\widehat{x}^{C'}\Sigma\widehat{x}^C\right]V_2 + x_{-1}'\Sigma x_{-1}V_3 + x_{-1}'\Sigma x^C V_4\right)}^{\text{Variability + Trading costs}}, \tag{3.14}$$

where x_{-1} is the investor's initial position, $x^C = \varsigma_1 x^M + \varsigma_2 x^{Min}$,

$$E\left[\widehat{x}^{C'}\Sigma\widehat{x}^C\right] = (c/\gamma^2)\left(\varsigma_1^2\left(\mu'\Sigma^{-1}\mu + (N/T)\right) + \varsigma_2^2\iota'\Sigma^{-1}\iota\right) + (c/\gamma^2)\left(2\varsigma_1\varsigma_2\mu'\Sigma^{-1}\iota\right), \tag{3.15}$$

and the $V_{i=2,3,4}$ account for the accumulated variability and trading costs:

$$V_1 = \frac{(1-\rho)(1-\beta)}{1-(1-\rho)(1-\beta)} \tag{3.16}$$

$$V_2 = \gamma\left(\frac{1-\rho}{\rho} + \frac{(1-\rho)(1-\beta)^2}{1-(1-\rho)(1-\beta)^2} - 2\frac{(1-\rho)(1-\beta)}{1-(1-\rho)(1-\beta)}\right) + \widetilde{\lambda}\frac{(1-\rho)\beta^2}{1-(1-\rho)(1-\beta)^2}, \tag{3.17}$$

$$V_3 = \gamma\frac{(1-\rho)(1-\beta)^2}{1-(1-\rho)(1-\beta)^2} + \widetilde{\lambda}\frac{(1-\rho)\beta^2}{1-(1-\rho)(1-\beta)^2}, \tag{3.18}$$

$$V_4 = 2\gamma\left(\frac{(1-\rho)(1-\beta)}{1-(1-\rho)(1-\beta)} - \frac{(1-\rho)(1-\beta)^2}{1-(1-\rho)(1-\beta)^2}\right) - 2\widetilde{\lambda}\frac{(1-\rho)\beta^2}{1-(1-\rho)(1-\beta)^2}. \tag{3.19}$$

From Proposition 8 we observe that as β goes to zero, V_2 and V_4 also approximate to zero. This implies that V_3 is the only element that defines the expected variability and trading costs of the multiperiod investor. Precisely, the investor's expected variability and trading costs are defined by $((1-\rho)/\rho)x_{-1}'\Sigma x_{-1}$, which is the accumulated variability of the investor's initial portfolio. Notice that when β is zero, trading costs do not affect the investor's expected utility.

In addition, we can observe that as the investor's initial position x_{-1} approximates to the static portfolio x^C, the expected return of the investor's initial portfolio in excess of the expected return of the static portfolio x^C, approximates to zero. Consequently, the optimal trading rate that we obtain from (3.14) must minimize the expected portfolio variability and trading costs.

To analyze the benefits of optimizing the trading rate, we study the relative loss for the multiperiod four-fund portfolio optimizing the trading rate as in (3.14), and the corresponding relative loss of the multiperiod four-fund portfolio with the nominal trading rate β as in (5). Figure 3.4 depicts the relative loss for our base-case investor with $\gamma = 10^{-8}$, $\lambda = 3 \times 10^{-7}$, $\rho = 1 - \exp(-0.1/260)$, and $T = 500$. As in the previous section, we define μ and Σ with the sample moments of the empirical dataset of commodity futures that we use in Section 3.5.

Figures 3.5a, 3.5b and 3.5c depict the relative loss for an investor whose initial portfolio is $x_{-1} = 0.1 \times x^M$ and we observe that the benefits from using the multiperiod four-fund portfolio that shrinks the trading rate are large. In particular, we observe that the investor can reduce the relative loss *more than a 15%.* Moreover, we observe that the relative loss of the different multiperiod portfolios remain almost invariant to changes in γ, λ and ρ. In addition, from Figure 3.5d we find that when the investor's initial portfolio is close to the static mean-variance portfolio, shrinking the trading rate β provides substantial benefits. In particular, when $x_{-1} \simeq 0.5 \times x^M$, one can reduce the relative loss to almost zero by shrinking the trading rate. In turn, shrinking the nominal trading rate may result into a considerable reduction of the investor's expected loss, specially in those situations where the investor's initial portfolio is close to the static mean-variance portfolio.

3.5. Out-of-sample performance evaluation

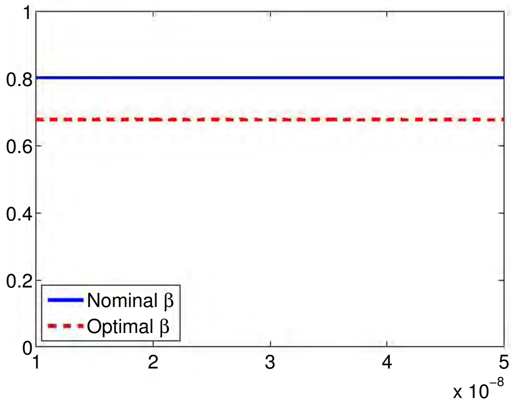
In this section, we compare the out-of-sample performance of the multiperiod shrinkage portfolios with that of the portfolios that ignore either transaction costs, parameter uncertainty, or both. We run the analysis with both simulated and empirical datasets.

3.5.1. Portfolio policies

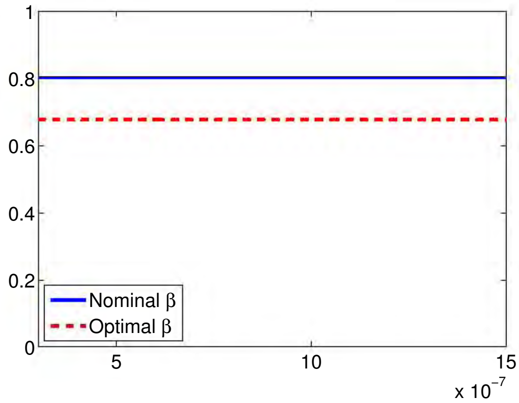
We consider seven different portfolio policies. We first consider three buy-and-hold portfolios based on single-period policies that ignore transaction costs. First, the sample Markowitz portfolio, which is the portfolio of an investor who ignores transaction costs and estimation error (S-M). Second, the single period two-fund shrinkage portfolio, which is the portfolio of an investor who ignores transaction costs, but takes into account estimation error by shrinking the Markowitz portfolio (S-2F). Specifically, this portfolio can be written as

$$x^{S2F} = \eta \widehat{x}^M, \tag{3.20}$$

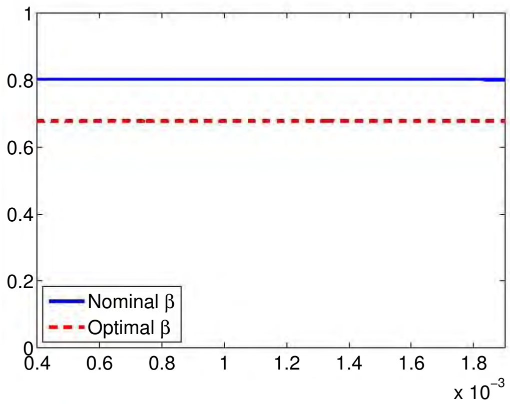
Figure 3.4.: Nominal Vs Optimal four-fund portfolios: Comparison of relative losses
This plot depicts the investor's relative loss for different values of γ, λ, and ρ. Our base-case investor is defined with $\gamma = 10^{-8}$, $\lambda = 3 \times 10^{-7}$ and $\rho = 1 - \exp(-0.1/260)$. We consider an investor that has 500 observations to construct the optimal trading strategy whose parameters are defined with the sample moments of the empirical dataset of commodity futures that we consider in the empirical application.



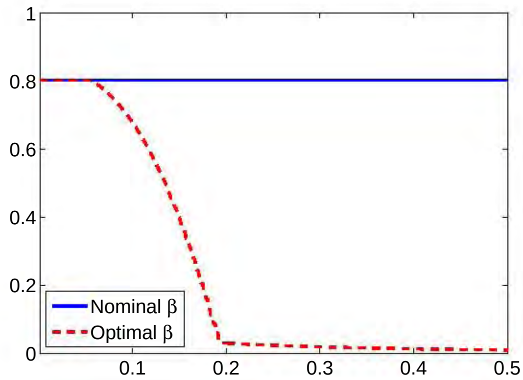
(a) Different values of γ



(b) Different values of λ



(c) Different values of ρ



(d) Different values of x_{-1}

where, as Kan and Zhou (2007) show, the optimal single-period shrinkage intensity η is as given by Proposition (7). The third portfolio is the single-period three-fund shrinkage portfolio of an investor who ignores transaction costs but takes into account estimation error by shrinking the Markowitz portfolio towards the minimum variance portfolio (S-3F-Min). Specifically, this portfolio can be written as

$$x^{S3F} = \varsigma_1 \widehat{x}^M + \varsigma_2 \widehat{x}^{Min}, \tag{3.21}$$

where the optimal single-period combination parameters are given in Proposition (7).

We then consider four multiperiod portfolios that take transaction costs into account. The first portfolio is the optimal portfolio policy of a multiperiod investor who takes into account transaction costs but ignores estimation error (M-M), which is given by Proposition 5. The second portfolio is the multiperiod three-fund shrinkage portfolio of an investor who shrinks the Markowitz portfolio (M3F), as given by Proposition (7). The third portfolio is the multiperiod four-fund shrinkage portfolio of an investor who combines the Markowitz portfolio with the minimum-variance portfolio (M4F-Min), as given by Proposition (7). The fourth portfolio is a modified version of the multiperiod four-fund shrinkage portfolio, where in addition the investor shrinks the trading rate by solving the optimization problem given by Proposition 8 (O-M4F-Min).

3.5.2. Evaluation methodology

We evaluate the out-of-sample portfolio gains for each strategy using a *rolling-window* approach similar to DeMiguel et al. (2009). To account for transaction costs in the empirical analysis, we define portfolio gains discounted by trading costs as:

$$\underline{r}^h_{l+1} = x^{h\prime}_l r_{l+1} - \widetilde{\lambda}\Delta x^{h\prime}_l \Sigma \Delta x^h_l, \tag{3.22}$$

where x^h_l denotes the estimated portfolio h at period l, r_l is the vector of price changes at time l, and Σ is the covariance matrix of asset prices.[9] Then, we compute the portfolio Sharpe ratio of all the considered trading strategies with the time series of the out-of-sample portfolio gains as:

$$SR^i = \frac{\overline{r}^h}{\sigma^h}, \tag{3.23}$$

$$\text{where } \ (\sigma^h)^2 = \frac{1}{L-T-1}\sum_{l=T}^{L-1}\left(x^{h\prime}_i r_{l+1} - \overline{r}^h\right)^2, \tag{3.24}$$

$$\overline{r}^h = \frac{1}{L-T}\sum_{l=T}^{L-1}\left(x^{h\prime}_l r_{l+1}\right), \tag{3.25}$$

where L is the total number of observations in the dataset, and T is the estimation window. We estimate the different portfolios using an estimation window of T=500 observations.[10]

We measure the statistical significance of the difference between the adjusted Sharpe ratios with the stationary bootstrap of Politis and Romano (1994) with B=1000 bootstrap samples and block size b=5.[11] Finally, we use the methodology suggested in (Ledoit and Wolf, 2008, Remark 2.1) to compute the resulting bootstrap p-values for the difference of every portfolio strategy with respect to the four-fund portfolio M4F-Min.

We consider an investor with a risk aversion parameter of $\gamma = 10^{-8}$, which corresponds with a relative risk aversion of one for a manager who has \$100M to trade. Garleanu and Pedersen (2012) consider an investor with a lower risk aversion parameter, but because our investor suffers from parameter uncertainty, it is reasonable to establish a higher risk aversion parameter. We use a discount factor ρ equal to $1 - \exp(-0.1/260)$, which corresponds with an annual discount of 10%. Finally, we consider transaction costs with $\lambda = 3 \times 10^{-7}$ as in Garleanu and Pedersen (2012). We subsequently test the robustness of our results to the values of these three parameters and observe that our main insights

[9]For the simulated data, we use the population covariance matrix, whereas for the empirical dataset with commodity futures we construct Σ with the sample estimate of the entire dataset.

[10]To compute those portfolios that account for parameter uncertainty, we need to estimate the optimal combination parameters, which require the true population moments. To mitigate the impact of parameter uncertainty in these parameters, we use the shrinkage vector of means proposed in DeMiguel et al. (2013), and the shrinkage covariance matrix by Ledoit and Wolf (2004b).

[11]We also compute the p-values when b=1, but we do not report these results to preserve space. These results are, however, equivalent to the block size b=5.

are robust.

Finally, we report the results for two different starting portfolios: the portfolio that is fully invested on the risk-free asset and the true Markowitz portfolio.[12] We have tried other starting portfolios such as the equally weighted portfolio and the portfolio that is invested in a single risky asset, but we observe that the results are similar and thus we do not report these cases to conserve space.

3.5.3. Simulated and empirical datasets

We first use simulation to generate two datasets with number of risky assets $N = 25$ and 50. The advantage of using simulated datasets is that they satisfy the assumptions underlying our analysis. Specifically, we simulate price changes from a multivariate normal distribution. We assume that the starting prices of all N risky assets are equal to one, and the annual average price changes are randomly distributed from a uniform distribution with support $[0.05, 0.12]$. In addition, the covariance matrix of asset price changes is diagonal with elements randomly drawn from a uniform distribution with support $[0.1, 0.5]$.[13] Without loss of generality, we set the return of the risk-free asset equal to zero. Under these specifications, a level of transaction costs of $\lambda = 3 \times 10^{-7}$ corresponds with a market that, on average, has a daily volume of $4.66 million.[14]

To understand the impact of data departing from the iid normal assumption, we consider an empirical dataset similar to that used by Garleanu and Pedersen (2012). Concretely, we construct a dataset with commodity futures of Aluminum, Copper, Nickel, Zinc, Lead, and Tin from the London Metal Exchange (LME), Gas Oil from the Intercontinental Exchange (ICE), WTI Crude, RBOB Unleaded Gasoline, and Natural Gas from the New York Mercantile Exchange (NYMEX), Gold and Silver from the New York Commodities Exchange (COMEX), and Coffee, Cocoa, and Sugar from the New York Board of Trade (NYBOT). We consider daily data from July 7th, 2004 until September 19th, 2012. We collect data from those commodity futures with 3-months maturity, and for those commodity futures where we do not find data with that contract specification (i.e. 3 months maturity), we collect the data of the commodity future with the largest time series. Some descriptive statistics and the contract multiplier for each commodity is provided in Table 3.1.[15]

3.5.4. Discussion of the out-of-sample performance

Table 3.2 reports the out-of-sample Sharpe ratios of the seven portfolio policies we consider on the three different datasets, together with the p-value of the difference between

[12]For the commodity dataset, we assume the true Markowitz portfolio is constructed with the entire sample.

[13]Notice that for our purpose of evaluating the impact of parameter uncertainty in an out-of-sample analysis, assuming that the covariance matrix is diagonal is not a strong assumption as we know that the investor's expected loss is proportional to $\theta = \mu' \Sigma^{-1} \mu$.

[14]To compute the trading volume of a set of assets worth 1$, we use the rule from Engle et al. (2012), where they assume that trading 1.59% of the daily volume implies a price change of 0.1%. Hence, for our first case we calculate the trading volume as $1.59\% \times \text{Trading Volume} \times 3 \times 10^{-7} \times 0.3^2 \times 0.5 = 0.1\%$.

[15]The contract multiplier specifies the number of units that are traded for each commodity in each contract. Also, notice that we do not report the trading volume. Unfortunately, we have not been able to obtain that type of data. However, we use the same level of transaction costs, which may be slightly high for the standard deviations of price changes that we have.

Table 3.1.: Commodity futures:

This table provides some descriptive statistics of the data from the commodity futures, as well as the contract multiplier.

Commodity	Average Price	Volatility price changes	Contract multiplier
Aluminium	56,231.71	888.37	25
Copper	161,099.45	3,268.96	25
Nickel	127,416.45	3,461.62	6
Zinc	54,238.84	1,361.69	25
Lead	45,925.04	1,227.02	25
Tin	78,164.60	1,733.53	5
Gasoil	69,061.48	1,571.89	100
WTI Crude	75,853.55	1,798.93	1000
RBOB Crude	88,503.62	2,780.74	42,000
Natural Gas	63,553.35	3,4439.78	10,000
Coffee	58,720.11	940.55	37,500
Cocoa	23,326.21	458.50	10
Sugar	18,121.58	462.35	112,000
Gold	94,780.87	1,327.11	100
Silver	87,025.94	2,415.69	5,000

the Sharpe ratio of every policy and that of the multiperiod four-fund shrinkage portfolio. Panels A and B give the results for a starting portfolio that is fully invested in the risk-free asset and a starting portfolio equal to the true Markowitz portfolio, respectively.

Comparing the multiperiod portfolios that take transaction costs into account with the static portfolios that ignore transaction costs, we find that the multiperiod portfolios substantially outperform the static portfolios. That is, we find that taking transaction costs into account has a substantial positive impact on performance.

Comparing the shrinkage portfolios with the portfolios that ignore transaction costs, we observe that shrinking helps both for the static and multiperiod portfolios. Specifically, we find that the portfolios that shrink only the Markowitz portfolio (S2F for the static case and M3F for the multiperiod case) outperform the equivalent portfolios that ignore estimation error (S-M for the static case and M-M for the multiperiod case). Moreover, we find that shrinking the Markowitz portfolio towards the minimum-variance portfolio improves performance substantially. Specifically, we observe that the S3F-Min and M4F-Min considerably outperform the shrinkage portfolios that shrink only the Markowitz portfolios (S2F and M3F).

Finally, our out-of-sample results confirm the insight from Section 3.4.2 that shrinking the trading rate may help when the starting portfolio is close to the true mean-variance portfolio. Specifically, we see from Panel A that shrinking the trading rate (in addition to shrinking the Markowitz portfolio towards the minimum-variance portfolio) does not result in any gains when the starting portfolio is fully invested in the risk-free asset, but Panel B shows that it may lead to substantial gains when the starting portfolio is the true mean-variance portfolio.

Overall, the best portfolio policy is the O-M4F-Min portfolio that shrinks the Markowitz portfolio towards the minimum-variance portfolio and, in addition, shrinks the trading

Table 3.2.: Sharpe ratio discounted with transaction costs

This table reports the annualized out-of-sample Share ratio for the different portfolio strategies that we consider. Sharpe ratios are discounted by quadratic transaction costs with $\lambda = 3 \times 10^{-7}$. The number in parentheses are the corresponding p-values for the difference of each portfolio strategy with the four-fund portfolio that combines the static mean-variance portfolio with the minimum-variance portfolio. Our considered base-case investor has an absolute risk aversion parameter of $\gamma = 10^{-8}$ and an impatience factor of $\rho = 1 - \exp(-0.1/260)$.

	Panel A: *Start from zero*			Panel B: *Start from* x^M		
	N=25	N=50	Com.	N=25	N=50	Com.
Static trading strategies						
S-M	-0.266	-0.345	-0.459	-0.266	-0.337	-0.452
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
S2F	0.076	0.105	0.102	0.068	0.101	0.106
	(0.000)	(0.000)	(0.050)	(0.000)	(0.000)	(0.036)
S3F-Min	0.678	0.633	0.739	0.678	0.637	0.769
	(0.000)	(0.000)	(0.126)	(0.000)	(0.000)	(0.148)
Multiperiod trading strategies						
M-M	0.150	0.297	0.056	0.153	0.295	0.052
	(0.000)	(0.008)	(0.036)	(0.004)	(0.008)	(0.042)
M3F	0.202	0.307	0.269	0.212	0.298	0.259
	(0.004)	(0.004)	(0.106)	(0.000)	(0.008)	(0.094)
M4F-Min	0.765	0.771	0.874	0.772	0.764	0.868
	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
O-M4F-Min	0.765	0.771	0.874	0.911	0.865	0.895
	(0.786)	(0.774)	(0.628)	(0.144)	(0.376)	(0.742)

rate while taking transaction costs into account. This portfolio policy outperforms the M4F-Min portfolio when the starting portfolio is close to the true minimum-variance portfolio, and it performs similar to the M4F-Min for other starting points. These two policies O-M4F-Min and M4F-Min appreciably outperform all other policies, which shows the importance of taking into account both transaction costs and estimation error.

We carry out an additional analysis to test the robustness of our results for different values of the risk-aversion parameter γ, and number of observations T. However, we do not report robustness checks for trading costs because only modifying parameter γ can provide equivalent results to those when we fix γ and modify λ.[16] We report these results in Table 3.3. We consider a base-case investor with an initial portfolio equal to the true Markowitz portfolio, $\gamma = 10^{-8}$, $\lambda = 3 \times 10^{-7}$, and $T = 500$.

[16]In particular, if we transform γ and λ by multiplying them with 10^{-z} and 10^{z}, respectively, we obtain the same multiperiod trading rate β, and in turn results are equivalent to those before the transformation. Then, if we want to study the impact of an increment/reduction on trading costs, we can simply reduce/increase γ by the same factor.

Table 3.3.: Sharpe ratio: some robustness checks (RC)

This table reports the annualized out-of-sample Share ratio for the different portfolio strategies that we consider. Our considered base-case investor has an absolute risk aversion parameter of $\gamma = 10^{-8}$ and an impatience factor of $\rho = 1 - \exp(-0.1/260)$ and faces quadratic transaction costs with $\lambda = 3 \times 10^{-7}$. The number in parentheses are the corresponding p-values for the difference of each portfolio strategy with the four-fund portfolio that combines the static mean-variance portfolio with the minimum-variance portfolio.

	Panel A: *RC for different* γ						Panel B: *RC for different* T					
	$\gamma = 10^{-9}$			$\gamma = 10^{-7}$			T=250			T=750		
	N=25	N=50	Com.	N=25	N=50	Com.	N=25	N=50	Com.	N=25	N=50	Com.
Static trading strategies												
S-M	-3.623	-4.020	-2.636	0.141	0.248	-0.044	-1.126	-1.458	-0.766	-0.023	0.207	-0.156
	(0.000)	(0.000)	(0.000)	(0.000)	(0.004)	(0.006)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.110)
S2F	-1.242	-1.383	-1.625	0.209	0.272	0.324	-0.186	-0.096	-0.073	0.087	0.435	-0.100
	(0.000)	(0.000)	(0.000)	(0.004)	(0.004)	(0.112)	(0.000)	(0.000)	(0.008)	(0.000)	(0.010)	(0.174)
S3F-Min	-0.195	-0.425	-0.740	0.765	0.748	0.984	0.349	0.416	0.089	0.718	0.822	0.417
	(0.000)	(0.000)	(0.000)	(0.066)	(0.076)	(0.030)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.900)
Multiperiod trading strategies												
M-M	0.086	0.219	-0.067	0.179	0.306	0.034	0.055	0.218	0.635	0.216	0.545	0.049
	(0.000)	(0.000)	(0.030)	(0.000)	(0.000)	(0.018)	(0.000)	(0.000)	(0.752)	(0.000)	(0.044)	(0.310)
M3F	0.194	0.305	0.115	0.226	0.290	0.311	0.171	0.321	0.515	0.179	0.569	-0.122
	(0.000)	(0.006)	(0.086)	(0.004)	(0.008)	(0.102)	(0.000)	(0.000)	(0.476)	(0.000)	(0.036)	(0.142)
M4F-Min	0.752	0.767	0.767	0.779	0.762	0.936	0.633	0.734	0.730	0.775	0.909	0.441
	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
O-M4F-Min	0.895	0.887	0.843	0.918	0.845	0.921	0.742	0.905	0.711	0.831	0.949	0.536
	(0.166)	(0.318)	(0.432)	(0.108)	(0.488)	(0.846)	(0.182)	(0.156)	(0.512)	(0.552)	(0.768)	(0.104)

In general, we observe that our main insights are robust to these parameters. There are substantial losses associated with ignoring both transaction costs and estimation error, and overall the best portfolio policies are M4F-Min and O-M4F-Min. We observe that for the simulated datasets shrinking the trading rate generally helps (that is, O-M4F-Min outperforms M4F-Min), although the difference between the Sharpe ratios of these two policies are not significant.

We also observe that the static portfolio policies are very sensitive to the risk-aversion parameter, and their performance is particularly poor for the case with low risk aversion γ. This is because investors with low risk aversion invest more on the risky assets and thus are more vulnerable to the impact of estimation error, which is particularly large for the static investors who ignore transaction costs. The multiperiod portfolio policies are more stable because taking transaction cots into account helps to combat estimation error, even for the case with low risk aversion. In particular, the difference of performance between static portfolios and multiperiod portfolios is large when the investor's risk aversion parameter is equal to $\gamma = 10^{-9}$.

Finally, we observe that the performance of the static portfolio strategies is also very sensitive to the choice of estimation window T. Specifically, static portfolios perform poorly when the estimation window is small and has $T = 250$ observations. For this estimation window, the difference between static mean-variance portfolios and multiperiod portfolios is large.

Summarizing, the out-of-sample losses associated with ignoring either transaction costs or parameter uncertainty are large. Moreover, overall the multiperiod four-fund shrinkage portfolio that combines the Markowitz portfolio with the minimum-variance portfolio achieves the best out-of-sample Sharpe ratio net of transaction costs. We also observe that shrinking the trading rate may provide considerable benefits, specially when the investor's initial portfolio is near the Markowitz portfolio.

3.6. Summary

We address the impact of parameter uncertainty in multiperiod portfolio selection with transation cots. We first provide a closed-form expression for the utility loss associated with using the plug-in approach to construct multiperiod portfolios. We observe from this closed-form expression that the investor's expected loss decreases with trading costs, the investor's impatience factor and the investor's risk aversion parameter.

Second, we propose a four-fund multiperiod shrinkage portfolio that mitigates the effects of estimation risk. We give closed-form expressions for the optimal shrinkage intensities, and we show that these intensities coincide with the shrinkage intensities for the corresponding single-period portfolio. In addition, we analytically characterize under which circumstances the four-fund shrinkage portfolio reduces the impact of parameter uncertainty, and we prove that it is prohibitive to use the plug-in multiperiod portfolio or the multiperiod shrinkage three-fund portfolio.

Third, we propose a novel technique that reduces the investor's trading rate to the static mean-variance portfolio, and we show that this methodology can substantially improve the investor's performance. In particular, we show that this methodology improves the investor's performance when the investor's initial position is close to the Markowitz portfolio.

Finally, our out-of-sample analysis with simulated and empirical datasets shows that the losses associated with ignoring transaction costs, parameter uncertainty, or both, are large, and that the four-fund shrinkage portfolio achieves good out-of-sample performance. In addition, we observe that shrinking the trading rate helps to mitigate the impact of parameter uncertainty and helps to attain high risk-adjusted expected returns.

4. Concluding remarks and future research

4.1. Conclusion

Parameter uncertainty is one of the main challenges of portfolio optimization. In particular, the classical mean-variance framework is very sensitive to small changes affecting the inputs of the model. The empirical evidence shows that small errors contaminating the inputs in the mean-variance framework result into large losses in the investor's expected performance. Therefore, parameter uncertainty must be considered by mean-variance investors to obtain portfolios that on average have larger risk-adjusted expected returns.

In this thesis, we contribute to the literature in two aspects. First, we study different calibration criteria for shrinkage estimators within the context of portfolio optimization. We consider shrinkage estimators for both the inputs of the mean-variance model – *shrinkage moments* – and the outputs – *shrinkage portfolios*. We provide analytical expressions for the optimal shrinkage intensity, and we also propose a novel nonparametric approach to compute the optimal shrinkage intensity. Finally, we evaluate the out-of-sample performance of the resulting portfolios with simulated and empirical datasets, and we find that the size of the shrinkage intensity plays a significant role on the investor's performance.

Second, we study the impact of parameter uncertainty in multiperiod portfolio selection with trading costs. Precisely, we characterize the expected loss of a multiperiod investor, and we find that it is equal to the product between the single-period utility loss in the absence of transaction costs, and another term that captures the multiperiod effects on the overall utility loss. In addition, we propose two multiperiod shrinkage portfolios to combat the impact of parameter uncertainty. In the first multiperiod shrinkage portfolio, we combine the Markowitz portfolio with the minimum-variance portfolio and we term as four-fund portfolio. This trading strategy results into an investment that diversifies the effects of estimation risk across the risk-free asset, the investor's current portfolio, the Markowitz portfolio and the minimum-variance portfolio. In the second multiperiod portfolio, we shrink the investor's trading rate. This novel technique limits the investor trading activity and it also helps to reduce the impact of parameter uncertainty. Finally, we characterize the out-of-sample performance of the proposed multiperiod shrinkage portfolios with simulated and empirical datasets, and we find that ignoring transaction costs, parameter uncertainty, or both, results into large losses.

As a result, investors must consider at least three elements in order to construct optimal portfolios. First, the investor should construct a portfolio in accordance with her own preferences; i.e. maximize her utility function. Second, the investor must take into consideration parameter uncertainty in order to understand the consequences of using sample estimates in portfolio optimization because estimated optimal portfolios may result into suboptimal solutions. Third, the investor should also consider frictions in the market because optimal solutions may be prohibitive in the presence of market constraints.

4.2. Future research lines

Overall, we address the inherent uncertainty that arises in portfolio optimization when the investor uses sample information to construct her optimal investment strategy. However, for tractability reasons we make some assumptions to characterize the effects of estimation risk in portfolio optimization. In particular, our main assumption is to consider that returns/price changes are independent and identically distributed (iid). Future research lines could relax this assumption and consider more general models that take into account serial dependence or predictive factors. Additionally, it could also be interesting to understand the impact of parameter uncertainty in financial markets as a whole, in contrast with the micro-perspective that we address in this thesis in which we characterize the impact of parameter uncertainty for a single investor.

These two research lines are natural extensions of this thesis. However, there are other research areas that have a direct implementation to portfolio optimization. For instance, modeling asset returns is an active area for econometricians and financial economists that has an immediate application to asset allocation.

In particular, a future research line would be modeling asset returns for large datasets where the number of assets may be larger than the number of observations. In this situation, we can use a factor model to capture the dynamics of asset returns. This is a parsimonious method that can model asset returns and mitigate the impact of parameter uncertainty, which is particularly of interest for large datasets. One possibility is to use those factors obtained from a principal component analysis (PCA). It is common in the literature of PCA to select those components that account for most of the variability in the dataset. However, it might be more interesting to look for those components that account for some other property such as an utility function. As a result, we could obtain components that model asset returns and have a clear financial interpretation based on the investor's utility function.

In general, models that explain better the dynamics of asset returns can help investors to construct optimal portfolios. One can apply the proposed model to define the statistical properties of asset returns and use these properties to characterize the investor's expected utility. Essentially, understanding asset returns gives a better idea of the dynamics of financial markets and this is indeed a worthwhile area to investigate further.

A. Proofs

A.1. Proof for Chapter 2

In this part, we prove all the propositions. Before going throughout all the propositions, we state two lemmas that will be used along the proofs:

Lemma 1. *Let x be a random vector in $\mathbb{R}^N$ with mean μ and covariance matrix Σ, and let A be a definite positive matrix in $\mathbb{R}^{N\times N}$. Thus, the expected value of the quadratic form $x'Ax$ is:*

$$E(x'Ax) = trace\,(A\Sigma) + \mu' A\mu. \tag{A.1}$$

The proof for the expected value of quadratic forms is a standard result in econometrics. See, for instance, (Greene, 2003, Page 49).

Lemma 2. *Given a sample $R \in \mathbb{R}^{T\times N}$ of independent and normally distributed observations, that is $R_t \sim N(\mu, \Sigma)$, the unbiased sample covariance matrix $\Sigma_{sp} = \frac{\sum_{t=1}^{T}\left(R_t - \overline{R}\right)^2}{T-1}$, where $\overline{R} = \frac{\sum_{t=1}^{T} R_t}{T}$, has a Wishart distribution $\Sigma_{sp} \sim \mathcal{W}\left(\frac{\Sigma}{T-1}, T-1\right)$. On the other hand, the unbiased estimator of the inverse covariance matrix $\Sigma_u^{-1} = \frac{T-N-2}{T-1}\Sigma_{sp}^{-1}$ has an inverse-Wishart distribution $\Sigma_u^{-1} \sim \mathcal{W}^{-1}\left((T-N-2)\Sigma^{-1}, T-1\right)$. Then, the expected values of $\Sigma_{sp}\Sigma_{sp}$, Σ_u^{-2} and $\Sigma_u^{-1}\Sigma\Sigma_u^{-1}$ are:*

$$E\left(\Sigma_{sp}\Sigma_{sp}\right) = \frac{T}{T-1}\Sigma^2 + \frac{1}{T-1} trace\,(\Sigma)\,\Sigma. \tag{A.2}$$

$$E\left(\Sigma_u^{-2}\right) = \frac{(T-N-2)}{(T-N-1)(T-N-4)}\left(trace\left(\Sigma^{-1}\right)\Sigma^{-1} + (T-N-2)\Sigma^{-2}\right), \tag{A.3}$$

$$E\left(\Sigma_u^{-1}\Sigma\Sigma_u^{-1}\right) = \frac{(T-N-2)(T-2)}{(T-N-1)(T-N-4)}\Sigma^{-1}. \tag{A.4}$$

The proof for $E\left(\Sigma_{sp}\Sigma_{sp}\right)$ can be found in Haff (1979), Theorem 3.1. The proof for $E\left(\Sigma_u^{-2}\right)$ and $E\left(\Sigma_u^{-1}\Sigma\Sigma_u^{-1}\right)$ are found in Haff (1979), Theorem 3.2.

A.1.1. Proof of Proposition 1

In this section, we prove the closed-form expression given in Proposition 1. In general, we consider that asset returns are independent and identically distributed. Then, from problem (2.6), we have:

$$\min_{\alpha} \quad E\left[\|\mu_{sh} - \mu\|_2^2\right] = (1-\alpha)^2 E\left[\|\mu_{sp} - \mu\|_2^2\right] + \alpha^2 \left\|\nu_\mu \iota - \mu\right\|_2^2. \tag{A.5}$$

Now, developing the optimality conditions of problem (A.5), we can obtain the optimal

α that minimizes the expected quadratic loss:

$$\alpha_\mu = \frac{E\left(\|\mu_{sp} - \mu\|_2^2\right)}{E\left(\|\mu_{sp} - \mu\|_2^2\right) + \|\nu_\mu \iota - \mu\|_2^2}, \tag{A.6}$$

where $\nu_\mu = \text{argmin}_\nu \|\nu\iota - \mu\|_2^2 = \overline{\mu}$. We develop the expected value given in (A.6) to derive the closed-form expression:

$$E\left(\|\mu_{sp} - \mu\|^2\right) = E\left(\mu_{sp}'\mu_{sp}\right) - \mu'\mu. \tag{A.7}$$

Since μ_{sp} is a random variable with mean μ and covariance matrix $\frac{\Sigma}{T}$, we can use Lemma 1 to obtain the closed-form expression of $E\left(\|\mu_{sp} - \mu\|^2\right)$. Thus:

$$E\left(\|\mu_{sp} - \mu\|^2\right) = (N/T)\overline{\sigma^2}, \tag{A.8}$$

where $\overline{\sigma^2} = \text{trace}\left(\Sigma\right)/N$, and it completes the proof.

A.1.2. Proof of Proposition 2

To prove this proposition we simply develop the optimality conditions from the calibration functions defined by the shrinkage portfolio formed with the sample and the target portfolios. The scale parameter is defined as $\nu = \text{argmin}\left\{\|\nu E(w_{tg}) - w_{op}\|_2^2\right\}$ with respect to ν. Developing the optimality conditions, we obtain that the optimal scale factor is $\nu = \frac{E(w_{tg})'w_{op}}{E(w_{tg})'E(w_{tg})}$. The expected quadratic loss function of the considered shrinkage portfolio is:

$$\begin{aligned} E\left(\|w_{sh} - w_{op}\|_2^2\right) &= E\left(\|(1-\alpha)\left(w_{sp} - w_{op}\right) + \alpha\left(\nu w_{tg} - w_{op}\right)\|_2^2\right) = \\ &= (1-\alpha)^2 E\left(\|w_{sp} - w_{op}\|_2^2\right) + \alpha^2 E\left(\|\nu w_{tg} - w_{op}\|_2^2\right) + \\ &+ 2(1-\alpha)\alpha E\left(\left(w_{sp} - w_{op}\right)'\left(\nu w_{tg} - w_{op}\right)\right). \end{aligned} \tag{A.9}$$

Therefore, developing the optimality conditions of $E\left(\|w_{sh} - w_{op}\|_2^2\right)$, we obtain that the optimal α is:

$$\alpha_{eql} = \frac{E\left(\|w_{sp} - w_{op}\|_2^2\right) - \tau_{sp-tg}}{E\left(\|w_{sp} - w_{op}\|_2^2\right) + E\left(\|\nu w_{tg} - w_{op}\|_2^2\right) - 2\tau_{sp-tg}}, \tag{A.10}$$

where $\tau_{sp-tg} = E\left(\left(w_{sp} - w_{op}\right)'\left(\nu w_{tg} - w_{op}\right)\right)$.

Second, the expected utility function of the shrinkage portfolio is:

$$\begin{aligned} E\left(U\left(w_{sh}\right)\right) &= (1-\alpha)E(w_{sp})'\mu + \alpha\nu E(w_{tg})'\mu - \\ &- \frac{\gamma}{2} E\left((1-\alpha)^2 w_{sp}'\Sigma w_{sp} + \alpha^2\nu^2 w_{tg}'\Sigma w_{tg} + 2(1-\alpha)\alpha\nu w_{sp}'\Sigma w_{tg}\right). \end{aligned} \tag{A.11}$$

Deriving the optimality conditions of the above expression, we obtain the optimal α:

$$\alpha_{ut} = \frac{E\left(w'_{sp}\Sigma w_{sp}\right) - \nu E\left(w'_{sp}\Sigma w_{tg}\right)}{E\left(w'_{sp}\Sigma w_{sp}\right) + \nu^2 E\left(w'_{tg}\Sigma w_{tg}\right) - 2\nu E\left(w'_{sp}\Sigma w_{tg}\right)} - \\ - \frac{1}{\gamma}\frac{E(w_{sp})'\mu - \nu E(w_{tg})'\mu}{E\left(w'_{sp}\Sigma w_{sp}\right) + \nu^2 E\left(w'_{tg}\Sigma w_{tg}\right) - 2\nu E\left(w'_{sp}\Sigma w_{tg}\right)}. \quad \text{(A.12)}$$

The proof of the variance is straightforward. The investor's portfolio variance is defined by the second addend of the utility, given by expression (A.11). Deriving the optimality conditions of that expression we have that the optimal α is:

$$\alpha_{var} = \frac{E\left(w'_{sp}\Sigma w_{sp}\right) - \nu E\left(w'_{sp}\Sigma w_{tg}\right)}{E\left(w'_{sp}\Sigma w_{sp}\right) + \nu^2 E\left(w'_{tg}\Sigma w_{tg}\right) - 2\nu E\left(w'_{sp}\Sigma w_{tg}\right)}. \quad \text{(A.13)}$$

A.1.3. Proof of Proposition 3

In this section, we prove the closed-form expressions of the expected values considered in Proposition 3. We consider that the vector of asset returns is iid normal. Thus, we develop the expected values of Proposition 3 and use Lemma 2 to derive the closed-form expressions:

$$E\left(\|\Sigma_{sp} - \Sigma\|_F^2\right) = \text{trace}\left(E\left(\Sigma'_{sp}\Sigma_{sp}\right) - \Sigma'\Sigma\right) = \frac{N}{T-1}\left(\frac{\text{trace}\left(\Sigma^2\right)}{N} + N\left(\overline{\sigma^2}\right)^2\right) \quad \text{(A.14)}$$

$$E\left(\left\|\Sigma_u^{-1} - \Sigma^{-1}\right\|_F^2\right) = \text{trace}\left(E\left(\Sigma_u^{-2}\right) - \Sigma^{-2}\right) = \text{trace}\left(\Omega\right) - \text{trace}\left(\Sigma^{-2}\right) \quad \text{(A.15)}$$

$$E\left(< \Sigma_u^{-1} - \Sigma^{-1}, \nu I - \Sigma^{-1} >\right) = \text{trace}\left(E\left(\Sigma_u^{-1} - \Sigma^{-1}\right)'\left(\nu I - \Sigma^{-1}\right)\right) = 0 \quad \text{(A.16)}$$

being $\overline{\sigma^2} = \text{trace}\left(\Sigma\right)/N$ and $\Omega = \frac{(T-N-2)}{(T-N-1)(T-N-4)}\left(\text{trace}\left(\Sigma^{-1}\right)\Sigma^{-1} + (T-N-2)\Sigma^{-2}\right)$. It completes the proof.

A.1.4. Proof of Proposition 4

Here, we illustrate how to prove Proposition 4. We develop each element mentioned in the Proposition. First, we show how to obtain $E\left(\left\|w_{sp}^{mv} - w_{op}^{mv}\right\|_2^2\right)$:

$$E\left(\left\|w_{sp}^{mv} - w_{op}^{mv}\right\|_2^2\right) = \frac{1}{\gamma^2}\left(E\left(\mu_{sp}\Sigma_u^{-2}\mu_{sp}\right) - \mu\Sigma^{-2}\mu\right). \quad \text{(A.17)}$$

Due to the fact that returns are assumed to be independent and normally distributed, μ_{sp} and Σ_{sp} are independent. Therefore, we can make use of Lemma 1 and Lemma 2 to compute the expected value of $E\left(\mu_{sp}\Sigma_u^{-2}\mu_{sp}\right)$. Thus:

$$E\left(\left\|w_{sp}^{mv}-w_{op}^{mv}\right\|_2^2\right)=\frac{1}{\gamma^2}\left[\frac{\text{trace}\left(\Sigma^{-1}\right)(T-N-2)(T-2)}{(T-N-1)(T-N-4)T}+\right.$$
$$+\frac{(T-N-2)}{(T-N-1)(T-N-4)}\left.\left[\text{trace}\left(\Sigma^{-1}\right)\mu'\Sigma^{-1}\mu+(T-N-2)\mu'\Sigma^{-2}\mu\right]\right]-$$
$$-\frac{1}{\gamma^2}\mu'\Sigma^{-2}\mu. \tag{A.18}$$

The following element is $E\left(\left\|\nu w_{sp}^{min}-w_{op}^{mv}\right\|_2^2\right)$:

$$E\left(\left\|\nu w_{sp}^{min}-w_{op}^{mv}\right\|_2^2\right)=\nu^2 E\left(\iota'\Sigma_u^{-2}\iota\right)+\frac{1}{\gamma^2}\mu'\Sigma^{-2}\mu-2\frac{\nu}{\gamma}\iota'\Sigma^{-2}\mu. \tag{A.19}$$

Using the value of $E\left(\Sigma_u^{-2}\right)$ given in Lemma 2, we have that:

$$E\left(\left\|\nu w_{sp}^{min}-w_{op}^{mv}\right\|_2^2\right)=\nu^2\frac{(T-N-2)}{(T-N-1)(T-N-4)}\left[\text{trace}\left(\Sigma^{-1}\right)\iota'\Sigma^{-1}\iota+\right.$$
$$\left.+(T-N-2)\iota'\Sigma^{-2}\iota\right]+\frac{1}{\gamma^2}\mu'\Sigma^{-2}\mu-2\frac{\nu}{\gamma}\iota'\Sigma^{-2}\mu. \tag{A.20}$$

Now, we prove how to obtain the closed-form expression of $E\left(\left\|w_{sp}^{min}-w_{op}^{min}\right\|_2^2\right)$. First, we expand the expression as usual:

$$E\left(\left\|w_{sp}^{min}-w_{op}^{min}\right\|_2^2\right)=E\left(\iota'\Sigma_u^{-2}\iota\right)-\iota'\Sigma^{-2}\iota. \tag{A.21}$$

Again, applying the value of $E\left(\Sigma_u^{-2}\right)$ given in Lemma 2, we obtain the following:

$$E\left(\left\|w_{sp}^{min}-w_{op}^{min}\right\|_2^2\right)=\frac{(T-N-2)}{(T-N-1)(T-N-4)}\left[\text{trace}\left(\Sigma^{-1}\right)\iota'\Sigma^{-1}\iota+\right.$$
$$\left.+(T-N-2)\iota'\Sigma^{-2}\iota\right]-\iota'\Sigma^{-2}\iota. \tag{A.22}$$

The remaining elements are easy to prove. Understanding how to apply Lemma1 and Lemma 2, expressions $E\left(w_{sp}^{mv\prime}\Sigma w_{sp}^{mv}\right)$, $E\left(w_{sp}^{min\prime}\Sigma w_{sp}^{min}\right)$ and $E\left(w_{sp}^{mv\prime}\Sigma w_{sp}^{min}\right)$ are simple to obtain. For instance,

$$E\left(w_{sp}^{mv\prime}\Sigma w_{sp}^{mv}\right)=\frac{1}{\gamma^2}E\left(\mu_{sp}\Sigma_u^{-1}\Sigma\Sigma_u^{-1}\mu_{sp}\right). \tag{A.23}$$

Since μ_{sp} and Σ_{sp} are independent, using Lemma 1 and the expression for $E\left(\Sigma_u^{-1}\Sigma\Sigma_u^{-1}\right)$ given in Lemma 2, we have:

$$E\left(w_{sp}^{mv\prime}\Sigma w_{sp}^{mv}\right)=\frac{1}{\gamma^2}\left(\frac{(T-N-2)(T-2)}{(T-N-1)(T-N-4)}\left(\frac{N}{T}+\mu'\Sigma^{-1}\mu\right)\right). \tag{A.24}$$

The proof of the remaining elements can be omitted since they are similar to the previous proof.

A.2. Proofs for Chapter 3

A.2.1. Proof Proposition 5

To solve the investor's problem, we first guess that the value function at any time i:

$$V(x_i) = -\frac{1}{2}x_i'Ax_i + x_i'B\mu + c. \tag{A.25}$$

Therefore, the Bellman equation becomes:

$$x_i'\mu - \frac{\gamma}{2}x_i'\Sigma x_i - \frac{\widetilde{\lambda}}{2}\Delta x_i'\Sigma\Delta x_i + (1-\rho)\left(-\frac{1}{2}x_i'Ax_i + x_i'B\mu + c\right), \tag{A.26}$$

where $\widetilde{\lambda} = (1-\rho)^{-1}\lambda$. The right hand side can be simplified as follows:

$$V(x_i) = -\frac{1}{2}x_i'Jx_i + x_i'h + l, \tag{A.27}$$

where $J = (\gamma+\widetilde{\lambda})\Sigma+(1-\rho)A$, $h = \mu+\widetilde{\lambda}\Sigma x_{i-1}+(1-\rho)B\mu$, and $l = -\frac{\widetilde{\lambda}}{2}x_{i-1}\Sigma x_{i-1}+(1-\rho)c$. The first-order necessary condition to solve the above problem give the optimal solution:

$$x_i = J^{-1}h. \tag{A.28}$$

Now, plugging the solution into the value function in (A.27), we obtain:

$$V^*(x_i) = \frac{1}{2}h'J^{-1}h + d. \tag{A.29}$$

From the above expression and using (A.25), we obtain that $A = -\widetilde{\lambda}^2\Sigma J^{-1}\Sigma + \widetilde{\lambda}\Sigma$ and $B = \widetilde{\lambda}\Sigma J^{-1}\left(I + (1-\rho)B\right)$. Thus, $A = \alpha\Sigma$, which implies that

$$\alpha = -\frac{\widetilde{\lambda}^2}{\gamma + \widetilde{\lambda} + (1-\rho)\alpha} + \widetilde{\lambda}. \tag{A.30}$$

Solving the above equation, we have that $\alpha = \frac{\sqrt{(\gamma+\widetilde{\lambda}\rho)^2+4\gamma\widetilde{\lambda}}-(\gamma+\widetilde{\lambda}\rho)}{2(1-\rho)}$. On the other hand, the solution for B is straightforward. It takes the form

$$B = \frac{\widetilde{\lambda}}{\gamma + \rho\widetilde{\lambda} + (1-\rho)\alpha}I. \tag{A.31}$$

Thus, the optimal solution, $x_i = J^{-1}h$, can be expressed as follows:

$$x_i = \frac{\widetilde{\lambda}}{\gamma + \widetilde{\lambda} + (1-\rho)\alpha}x_{i-1} + \frac{\gamma + (1-\rho)\gamma B}{\gamma + \widetilde{\lambda} + (1-\rho)\alpha}\frac{1}{\gamma}\Sigma\mu. \tag{A.32}$$

The above expression can be simplified as follows (see Garleanu and Pedersen (2012)):

$$x_i = (1-\beta)\,x_{i-1} + \beta\frac{1}{\gamma}\Sigma\mu. \tag{A.33}$$

where $\beta = \alpha/\widetilde{\lambda}$.
To prove the monotonicity of the convergence rate β, we only need to analyze the derivative of β with respect to γ, λ and ρ.

First, we show that the convergence rate β is a monotonic and nondecreasing function with respect to γ. Thus, we show that the derivative of β with respect to γ is always positive for any γ, λ, $\rho \geq 0$. To do that, it suffices to show that

$$2 \times (1-\rho) \times \frac{\partial \alpha}{\partial \gamma} \geq 0.$$

Then,

$$2 \times (1-\rho) \times \frac{\partial \alpha}{\partial \gamma} = \frac{1}{2} \frac{2\left(\gamma + \widetilde{\lambda}\rho\right) + 4\lambda}{\sqrt{\left(\gamma + \widetilde{\lambda}\rho\right)^2 + 4\gamma\lambda}} - 1 > 0 \Rightarrow \left(\gamma + \widetilde{\lambda}\rho\right) + 2\lambda \geq \sqrt{\left(\gamma + \widetilde{\lambda}\rho\right)^2 + 4\gamma\lambda}. \tag{A.34}$$

Now, we take the square of the above inequality, which is a monotone transformation and does not affect the results. Then:

$$\left(\gamma + \widetilde{\lambda}\rho\right)^2 + 4\lambda^2 + 4\lambda\left(\gamma + \widetilde{\lambda}\rho\right) \geq \left(\gamma + \widetilde{\lambda}\rho\right)^2 + 4\gamma\lambda \Rightarrow \tag{A.35}$$

$$\Rightarrow 4\lambda^2 + 4\gamma\lambda + 4\lambda\widetilde{\lambda}\rho \geq 4\gamma\lambda. \tag{A.36}$$

Inequality (A.36) is always true for any γ, λ, $\rho \geq 0$.

To prove that the rate of convergence β is a monotonic decreasing function with respect to λ, we show that the derivative of β with respect to λ is negative. First, let us define $\phi = \rho/(1-\rho)$. Thus,

$$2 \times \frac{\left(\frac{(\gamma+\lambda\phi)\phi+2\gamma}{\sqrt{(\gamma+\lambda\phi)^2+4\gamma\lambda}} - \phi\right)\lambda - \left(\sqrt{(\gamma+\lambda\phi)^2 + 4\gamma\lambda} - (\gamma+\lambda\phi)\right)}{4\lambda^2} < 0. \tag{A.37}$$

To prove that the above inequality holds, it suffices to prove that the numerator is negative. Thus,

$$\left(\frac{1}{2}\frac{2(\gamma+\lambda\phi)\phi + 4\gamma}{\sqrt{(\gamma+\lambda\phi)^2 + 4\gamma\lambda}} - \phi\right)\lambda - \left(\sqrt{(\gamma+\lambda\phi)^2 + 4\gamma\lambda} - (\gamma+\lambda\phi)\right) < 0. \tag{A.38}$$

After some straightforward manipulations, we have that

$$\left((\gamma+\lambda\phi)\phi + 2\gamma\right)\lambda < (\gamma+\lambda\phi)^2 + 4\gamma\lambda - \sqrt{(\gamma+\lambda\phi)^2 + 4\gamma\lambda} \times \gamma. \tag{A.39}$$

The above inequality can be expressed as:

$$\gamma\phi\lambda + \lambda^2\phi^2 + 2\gamma\lambda < \gamma^2 + \lambda^2\phi^2 + 2\gamma\phi\lambda + 4\gamma\lambda - \sqrt{(\gamma+\gamma\phi)^2 + 4\gamma\lambda} \times \gamma, \tag{A.40}$$

which may be simplified as

$$0 < \gamma^2 + \gamma\phi\lambda + 2\gamma\lambda - \sqrt{(\gamma+\gamma\phi)^2 + 4\gamma\lambda} \times \gamma. \tag{A.41}$$

Dividing by γ, and taking the square, we have:

$$(\gamma + \lambda\phi)^2 + 4\gamma\lambda < (\gamma + \lambda\phi)^2 + 4\lambda^2 + 4(\gamma + \lambda\phi)\lambda \Rightarrow \tag{A.42}$$
$$\Rightarrow 0 < 4\lambda^2 + 4\lambda^2\phi, \tag{A.43}$$

which shows that for any γ, λ, $\rho > 0$, the rate of convergence β is a monotonic decreasing function with respect to λ.

Finally, to prove that the rate of convergence β is a monotonic decreasing function with respect to ρ, we show that:

$$2 \times \lambda \times \frac{\partial \beta}{\partial \rho} = \frac{1}{2} \frac{2\left(\gamma + \lambda\frac{\rho}{1-\rho}\right)\frac{\lambda}{(1-\rho)^2}}{\sqrt{\left(\gamma + \lambda\frac{\rho}{1-\rho}\right)^2 + 4\gamma\lambda}} - \frac{\lambda}{(1-\rho)^2} < 0. \tag{A.44}$$

After some straightforward manipulations, we have that

$$\left(\gamma + \lambda\frac{\rho}{1-\rho}\right) < \sqrt{\left(\gamma + \lambda\frac{\rho}{1-\rho}\right)^2 + 4\gamma\lambda}. \tag{A.45}$$

Now, taking the square of the above inequality, we have:

$$\left(\gamma + \lambda\frac{\rho}{1-\rho}\right)^2 < \left(\gamma + \lambda\frac{\rho}{1-\rho}\right)^2 + 4\gamma\lambda, \tag{A.46}$$

which holds for any γ, λ, $\rho > 0$, and thus it completes the proof that ensures that the rate of convergence β is a monotonic decreasing function with respect to ρ.

A.2.2. Proof of Proposition 6

To prove Proposition 6, we first write the investor's expected loss:

$$\begin{aligned}\delta(\{x_i\}, \{\widehat{x}_i\}) = \sum_{i=0}^{\infty}(1-\rho)^{i+1}\Bigg\{ & x_i'\mu - \frac{\gamma}{2}x_i'\Sigma x_i - \frac{\widetilde{\lambda}}{2}\Delta x_i'\Sigma\Delta x_i \\ & -E\left[\widehat{x}_i'\mu - \frac{\gamma}{2}\widehat{x}_i'\Sigma\widehat{x}_i - \frac{\widetilde{\lambda}}{2}\Delta\widehat{x}_i'\Sigma\Delta\widehat{x}_i\right]\Bigg\}.\end{aligned} \tag{A.47}$$

And from the above expression, it is easy to see that the investor's expected loss is:

$$\delta(\{x_i\}, \{\widehat{x}_i\}) = \sum_{i=0}^{\infty}(1-\rho)^{i+1}\left\{E\left[\frac{\gamma}{2}\widehat{x}_i'\Sigma\widehat{x}_i + \frac{\widetilde{\lambda}}{2}\Delta\widehat{x}_i'\Sigma\Delta\widehat{x}_i\right] - \frac{\gamma}{2}x_i'\Sigma x_i - \frac{\widetilde{\lambda}}{2}\Delta x_i'\Sigma\Delta x_i\right\}. \tag{A.48}$$

Now, we can plug the estimated investor's optimal strategy in (A.48) to obtain a simplified expression of the investor's expected loss. Moreover, all those elements that are linear functions with respect to the sample Markowitz portfolio disappear due to the unbiasedness of the estimator. Then, we use the following expression for the estimated multiperiod portfolio:

$$\widehat{x}_i = (1-\beta)^{i+1}x_{-1} + \beta\xi_i\widehat{x}^M \text{ and } \Delta\widehat{x}_i = \phi_i x_{-1} + \beta(1-\beta)^i\widehat{x}^M, \tag{A.49}$$

where $\xi_i = \sum_{j=0}^{i}(1-\beta)^j$ and $\phi = ((1-\beta)^{i+1} - (1-\beta)^i)$. Then, after some straightforward manipulations, we obtain that the investor's expected loss is:

$$\delta(\{x_i\},\{\widehat{x}_i\}) = \frac{1}{2\gamma}\left(E\left[\widehat{\mu}'\widehat{\Sigma}^{-1}\Sigma\widehat{\Sigma}^{-1}\widehat{\mu}\right] - \theta\right) \times \sum_{i=0}^{\infty}(1-\rho)^{i+1}\left[AV_i + AC_i\right], \tag{A.50}$$

where $\theta = \mu'\Sigma^{-1}\mu$, $AV_i = \beta^2\xi_i^2$ stands for the accumulated portfolio variability and $AC_i = \beta^2(\widetilde{\lambda}/\gamma)(1-\beta)^{2i}$ stands for the accumulated trading costs. Then, we can substitute $(1/2\gamma)(E\left[\widehat{\mu}'\widehat{\Sigma}^{-1}\Sigma\widehat{\Sigma}^{-1}\widehat{\mu}\right] - \theta)$ with $\delta(x^M, \widehat{x}^M)$, and make the following simplifications for geometric series:

$$\xi_i = \sum_{j=0}^{i}(1-\beta)^j = \frac{1-(1-\beta)^{i+1}}{\beta}. \tag{A.51}$$

In turn, we obtain that

$$\beta^2\sum_{i=0}^{\infty}(1-\rho)^{i+1}\xi_i^2 = \sum_{i=0}^{\infty}(1-\rho)^{i+1} + \sum_{i=0}^{\infty}(1-\rho)^{i+1}\left[(1-\beta)^{2i+2} - 2(1-\beta)^{i+1}\right]. \tag{A.52}$$

Because $(1-\rho)$ and $(1-\beta)$ are positive elements and smaller than one, we can express the above geometric series as follows:

$$AV = \beta^2\sum_{i=0}^{\infty}(1-\rho)^{i+1}\xi_i^2 = \frac{1-\rho}{\rho} + \frac{(1-\rho)(1-\beta)^2}{1-(1-\rho)(1-\beta)^2} - 2\frac{(1-\rho)(1-\beta)}{1-(1-\rho)(1-\beta)}. \tag{A.53}$$

Now, applying the same arguments, we can simplify the following expression:

$$AC = \sum_{i=0}^{\infty}(1-\rho)^{i+1}\beta^2\frac{\widetilde{\lambda}}{\gamma}(1-\beta)^{2i} = \frac{\lambda}{\gamma}\frac{\beta^2}{1-(1-\rho)(1-\beta)^2}. \tag{A.54}$$

In turn, we obtain that the investor's expected loss is

$$\delta(\{x_i\},\{\widehat{x}_i\}) = \delta(x^M, \widehat{x}^M) \times [AV + AC]. \tag{A.55}$$

A.2.3. Proof of Proposition 7

We now prove that the optimal combination parameter of multiperiod portfolios coincide with the optimal combination parameter in the static framework. First, let us define the investor's initial portfolio as x_{-1}. Then, we can write the investor's four-fund portfolio as:

$$\widehat{x}_i = (1-\beta)^{i+1}x_{-1} + \beta\xi_i\widehat{x}^C, \tag{A.56}$$

where $\widehat{x}^C = (\varsigma_1 \widehat{x}^M + \varsigma_2 \widehat{x}^{Min})$, and

$$\Delta \widehat{x}_i = \phi_i x_{-1} + \beta(1-\beta)^i \widehat{x}^C, \tag{A.57}$$

where $\xi_i = \sum_{j=0}^{i}(1-\beta)^j$ and $\phi = ((1-\beta)^{i+1} - (1-\beta)^i)$. Then, the investor's expected utility is defined as:

$$\begin{aligned} E\Bigg[\sum_{i=0}^{\infty}(1-\rho)^{i+1}\Big\{&(1-\beta)^{i+1}x'_{-1}\mu + \beta\xi_i \widehat{x}^{C'}\mu \\ &- \frac{\gamma}{2}\Big((1-\beta)^{2i}x'_{-1}\Sigma x_{-1} + \beta^2\xi_i^2\widehat{x}^{C'}\Sigma\widehat{x}^C + 2(1-\beta)^{i+1}\xi_i x'_{-1}\Sigma\widehat{x}^C\Big) \\ &- \frac{\widetilde{\lambda}}{2}\Big(\phi_i^2 x'_{-1}\Sigma x_{-1} + \beta^2(1-\beta)^{2i}\widehat{x}^{C'}\Sigma\widehat{x}^C + 2\phi_i\beta(1-\beta)^i x'_{-1}\Sigma\widehat{x}^C\Big)\Big\}\Bigg] \end{aligned} \tag{A.58}$$

The above expression can be simplified with the following properties of geometric series:

$$\beta\xi_i = \beta\frac{1-(1-\beta)^{i+1}}{\beta} = 1-(1-\beta)^{i+1} \tag{A.59}$$

$$\beta^2\xi_i^2 = 1 + (1-\beta)^{2+2} - 2(1-\beta)^{i+1} \tag{A.60}$$

$$r_1 = \sum_{i=0}^{\infty}(1-\rho)^{i+1}(1-\beta)^{i+1} = \frac{(1-\rho)(1-\beta)}{1-(1-\rho)(1-\beta)} \tag{A.61}$$

$$r_2 = \sum_{i=0}^{\infty}(1-\rho)^{i+1}(1-\beta)^{2i+2} = \frac{(1-\rho)(1-\beta)^2}{1-(1-\rho)(1-\beta)^2} \tag{A.62}$$

$$r_3 = \sum_{i=0}^{\infty}(1-\rho)^{i+1}(1-\beta)^{2i} = \frac{(1-\rho)}{1-(1-\rho)(1-\beta)^2} \tag{A.63}$$

$$r_4 = \sum_{i=0}^{\infty}(1-\rho)^{i+1}(1-\beta)^{2i+1} = \frac{(1-\rho)(1-\beta)}{1-(1-\rho)(1-\beta)^2} \tag{A.64}$$

$$r_5 = \sum_{i=0}^{\infty}(1-\rho)^{i+1} = \frac{(1-\rho)}{\rho} \tag{A.65}$$

And in turn, the investor's expected utility can be simplified as follows:

$$\begin{aligned} &r_1(x_{-1}-x^C)'\mu + \frac{1-\rho}{\rho}x^{C'}\mu - \frac{\gamma}{2}\Big\{r_2 x'_{-1}\Sigma x_{-1} + (r_5+r_2-2r_1)\,E(\widehat{x}^{C'}\Sigma\widehat{x}^C) \\ &+2x_{-1}\Sigma x^C(r_1-r_2)\Big\} - \frac{\widetilde{\lambda}}{2}\Big\{\beta^2 r_3 x'_{-1}\Sigma x_{-1} + E(\widehat{x}^{C'}\Sigma\widehat{x}^C)\beta^2 r_3 + 2\beta(r_4-r_3)x'_{-1}\Sigma x^C\Big\} \end{aligned} \tag{A.66}$$

Now, we develop the first order conditions with respect to ς_1, and we obtain that the optimal value is:

$$\varsigma_1 = \frac{E\left[\widehat{x}^{M'}\mu\right]}{\gamma E\left[\widehat{x}^{M'}\Sigma\widehat{x}^M\right]}\frac{W_1}{W_2} - \frac{x'_{-1}\Sigma x^M}{\gamma E\left[\widehat{x}^{M'}\Sigma\widehat{x}^M\right]}\frac{W_3}{W_2} - \varsigma_2\frac{E\left[\widehat{x}^{M'}\Sigma\widehat{x}^{Min}\right]}{E\left[\widehat{x}^{M'}\Sigma\widehat{x}^M\right]}, \tag{A.67}$$

where $W_1 = r_5 - r_1$, $W_2 = (r_5 + r_2 - 2r_1) + (\widetilde{\lambda}/\gamma)\beta^2 r_3$, and $W_3 = \gamma(r_1 - r_2) + \widetilde{\lambda}\beta(r_4 - r_3)$. We numerically verify that $W_1/W_2 = 1$ and $W_3 = 0$, so that the optimal parameter ς_1

takes the following expression:

$$\varsigma_1 = \frac{E\left[\widehat{x}^{M\prime}\mu\right]}{\gamma E\left[\widehat{x}^{M\prime}\Sigma\widehat{x}^{M}\right]} - \varsigma_2 \frac{E\left[\widehat{x}^{M\prime}\Sigma\widehat{x}^{Min}\right]}{E\left[\widehat{x}^{M\prime}\Sigma\widehat{x}^{M}\right]}. \tag{A.68}$$

Accordingly, the optimal value of ς_2 is

$$\varsigma_2 = \frac{E\left[\widehat{x}^{Min\prime}\mu\right]}{\gamma E\left[\widehat{x}^{Min\prime}\Sigma\widehat{x}^{Min}\right]} - \varsigma_1 \frac{E\left[\widehat{x}^{M\prime}\Sigma\widehat{x}^{Min}\right]}{E\left[\widehat{x}^{Min\prime}\Sigma\widehat{x}^{Min}\right]}. \tag{A.69}$$

Therefore, one can solve the system given by (A.68)-(A.69) to obtain the optimal values of ς_1 and ς_2. This corresponds with the system of linear equations that one has to solve to obtain the optimal combination parameters in the static framework. In turn, we obtain; see Kan and Zhou (2007):

$$\varsigma_1 = c^{-1}\frac{\Psi^2}{\Psi^2 + \frac{N}{T}}, \tag{A.70}$$

$$\varsigma_2 = c^{-1}\frac{\frac{N}{T}}{\Psi^2 + \frac{N}{T}} \times \frac{\mu'\Sigma^{-1}\iota}{\iota'\Sigma^{-1}\iota}, \tag{A.71}$$

where $c = [(T-2)(T-N-2)]/[(T-N-1)(T-N-4)]$ and $\Psi^2 = \mu'\Sigma^{-1}\iota - (\mu'\Sigma^{-1}\iota)^2/(\iota'\Sigma^{-1}\iota) > 0$. Accordingly, one can obtain the optimal value of η by setting $\varsigma_2 = 0$ in equation (A.68), and we obtain that the optimal value of η is:

$$\eta = \frac{E\left[\widehat{x}^{M\prime}\mu\right]}{\gamma E\left[\widehat{x}^{M\prime}\Sigma\widehat{x}^{M}\right]} = c^{-1}\frac{\mu'\Sigma^{-1}\mu}{\mu'\Sigma^{-1}\mu} = c^{-1}. \tag{A.72}$$

A.2.4. Proof of Corollary 1

We know from Proposition 7 that the optimal combination parameters coincide with the optimal combination parameters of the static case. Then, we can show that it is optimal to shrink the static mean-variance portfolio if the derivative of the investor's (static) expected utility with respect to parameter η is negative when $\eta = 1$. Deriving the investor's expected utility with respect to η and setting $\eta = 1$, we obtain the it is optimal to have $\eta < 1$ when:

$$E\left(\widehat{x}^{M\prime}\mu\right) < \gamma E\left(\widehat{x}^{M\prime}\Sigma\widehat{x}^{M}\right). \tag{A.73}$$

If we characterize the expectations from the above expression, we obtain that $\eta < 1$ if $1 < c$, where $c = [(T-N-2)(T-2)]/[(T-N-1)(T-N-4)]$. Because, $c > 1$, we observe that it is always optimal to shrink the static mean-variance portfolio.

Now, if we take derivatives of the investor's (static) expected utility with respect to parameter ς_2, and then set $\varsigma_2 = 0$, this derivative if positive (an in turn it is optimal to have $\varsigma_2 > 0$) if

$$E\left(\widehat{x}^{T\prime}\mu\right) > \gamma\varsigma_1 E\left(\widehat{x}^{M\prime}\Sigma\widehat{x}^{T}\right). \tag{A.74}$$

Now, characterizing the above expectations, we obtain that $\varsigma_2 > 0$ if $1 > \varsigma_1 c$. From the optimal expression of ς_1, we obtain that $1 > \varsigma_1 c$ if $1 > \Psi^2/(\Psi^2 + N/T)$, which always holds because Ψ^2 can be written as $\Psi^2 = (\mu - \mu_g)'\Sigma^{-1}(\mu - \mu_g)$, where $\mu_g = (\iota'\Sigma^{-1}\mu)/(\iota'\Sigma^{-1}\iota)$, and in turn Ψ^2 is nonnegative. Moreover, from the optimal expression for ς_2, we observe that the optimal value is always positive because $\Psi^2 = \mu'\Sigma^{-1}\iota - (\mu'\Sigma^{-1}\iota)^2/(\iota'\Sigma^{-1}\iota) > 0$, and it means that $\mu'\Sigma^{-1}\iota$ should be positive, otherwise $\Psi^2 > 0$ would not hold. this means that all the elements require to compute the optimal ς_2 are positive, and in turn the optimal ς_2 is positive.

A.2.5. Proof of Proposition 8

Writing the expected utility for an investor using the four-fund portfolio as in (A.66), it is straightforward to see that we can obtain the optimal β that minimizes the investor's expected loss by solving the following problem:

$$V_1(x_{-1} - x^C)'\mu - \frac{1}{2}\left(E\left[\widehat{x}^{C'}\Sigma\widehat{x}^C\right]V_2 + x'_{-1}\Sigma x_{-1}V_3 + x'_{-1}\Sigma x^C V_4\right), \tag{A.75}$$

where V_i accounts for the accumulated variability and trading costs of $\widehat{x}^C$ and the investor's initial position x_{-1}, and they take the form:

$$V_1 = \frac{(1-\rho)(1-\beta)}{1-(1-\rho)(1-\beta)} \tag{A.76}$$

$$V_2 = \gamma\left(\frac{(1-\rho)}{\rho} + \frac{(1-\rho)(1-\beta)^2}{1-(1-\rho)(1-\beta)^2} - 2\frac{(1-\rho)(1-\beta)}{1-(1-\rho)(1-\beta)}\right) + \widetilde{\lambda}\frac{(1-\rho)\beta^2}{1-(1-\rho)(1-\beta)^2}, \tag{A.77}$$

$$V_3 = \gamma\frac{(1-\rho)(1-\beta)^2}{1-(1-\rho)(1-\beta)^2} + \widetilde{\lambda}\frac{(1-\rho)\beta^2}{1-(1-\rho)(1-\beta)^2}, \tag{A.78}$$

$$V_4 = 2\gamma\left(\frac{(1-\rho)(1-\beta)}{1-(1-\rho)(1-\beta)} - \frac{(1-\rho)(1-\beta)^2}{1-(1-\rho)(1-\beta)^2}\right) - 2\widetilde{\lambda}\frac{(1-\rho)\beta^2}{1-(1-\rho)(1-\beta)^2}. \tag{A.79}$$

Now, we characterize $E\left[\widehat{x}^{C'}\Sigma\widehat{x}^C\right]$, which is defined as:

$$E\left[\widehat{x}^{C'}\Sigma\widehat{x}^C\right] = \frac{c}{\gamma^2}\left(\varsigma_1^2\left(\mu'\Sigma^{-1}\mu + \frac{N}{T}\right) + \varsigma_2^2\iota'\Sigma^{-1}\iota + 2\varsigma_1\varsigma_2\mu'\Sigma^{-1}\iota\right), \tag{A.80}$$

where $c = [(T-N-2)(T-2)]/[(T-N-1)(T-N-4)$.

B. Robustness checks for Chapter 2

B.1. Overview

In this supplementary appendix, we report four different sets of additional empirical results that demonstrate the robustness of our analysis. First, we report the out-of-sample turnover for all the portfolios that we consider in Chapter 2. Second, we report the results for a longer estimation window with M=150 observations. Third, we report the results for a smaller estimation window with M=60 observations. For the case with M=60, we only consider minimum variance portfolios that shrink the covariance matrix, because the rest of the approaches result in a singular (or nearly singular) matrix for the 48IndP, 100FF and SP100 datasets. Fourth, we compare the out-of-sample performance of the minimum-variance portfolios constructed with the shrinkage covariance matrices calibrated with the approach proposed in Section 2.1.4, which accounts for the RIAL and the condition number, and with shrinkage covariance matrices calibrated with a related simpler approach that we term the direct approach.

In general, the results that we highlight in Chapter 2 are very robust. First, the considered calibration criterion for the shrinkage vector of means provides more stable mean-variance portfolios with higher net Sharpe ratio than those constructed with the shrinkage vector of means of Jorion (1986). Second, the condition number of the covariance matrix matters to calibrate its shrinkage estimator and it helps to obtain more stable portfolios with larger net Sharpe ratio. Third, for those shrinkage portfolios that consider the vector of means, the variance criterion provides in general the best results, whereas for the shrinkage portfolio that does not consider the vector of means the most robust criterion is the expected quadratic loss minimization criterion. Finally, we observe that the proposed multivariate smoothed bootstrap effectively captures the departure of data from normality and it helps to provide better out-of-sample results.

In the next four sections we report the results for the four different sets of additional experiments.

B.2. Out-of-sample Turnover for M=120

In this part, we report the out-of-sample turnover of all the studied portfolios. We do not include them in the main body of the chapter because they are implicitly considered in the Sharpe ratio of returns net of transaction costs. The definition of the out-of-sample turnover is as follows:

$$\text{Turnover}^i = \frac{1}{T-M-1}\sum_{t=M}^{T-1}\sum_{j=1}^{N}\left(\left|w^i_{j,t+1} - w^i_{j,t^+}\right|\right), \tag{B.1}$$

where $w^i_{j,t}$ denotes the estimated portfolio weight of asset j at time t under policy i and w^i_{j,t^+} is the estimated portfolio weight of asset j accumulated at time $t+1$, which implies that the turnover is equal to the sum of the absolute value of the rebalancing trades across the N available assets over the $T - M - 1$ trading dates, normalized by the total number of trading dates.

In general, we observe from Tables B.1 and B.2 that the turnover results are consistent with what we report in Chapter 2. First, the criterion for the shrinkage vector of means provides more stable mean-variance portfolios with lower turnover than the benchmark mean-variance portfolios. Second, the matrix condition number matters and helps to provide more stable minimum-variance portfolios when using a shrinkage covariance matrix. The variance and expected quadratic loss criteria provide more stable portfolios with lower turnover for the shrinkage portfolios that consider the vector of means and the shrinkage portfolio that does not consider the vector of means, respectively. Finally, we observe that the studied smoothed bootstrap also provides estimated portfolios with lower turnover.

Table B.1.: Turnover of benchmark portfolios and portfolios estimated with shrinkage moments

This table reports the out-of-sample turnover of benchmark portfolios and portfolios constructed by using the studied shrinkage estimators for the moments of asset returns. We consider an investor with a risk aversion parameter of $\gamma = 5$.

Policy	5IndP	10IndP	38IndP	48IndP	100FF	SP100
Panel A: Benchmark Portfolios						
Portfolios that consider the vector of means						
mv	0.372	0.482	1.854	3.685	277.548	5.415
bs	0.192	0.255	1.051	2.103	393.893	4.492
Portfolios that do not consider the vector of means						
min	0.094	0.153	0.495	0.793	6.639	2.730
lw	0.065	0.104	0.274	0.377	0.904	0.302
lw-m	0.093	0.134	0.308	0.371	1.063	0.266
Naïve Portfolios						
1/N	0.018	0.024	0.031	0.033	0.023	0.062
Panel B: Portfolios calibrated parametrically						
Portfolios that do not consider the vector of means						
par-lw	0.074	0.119	0.339	0.480	1.398	0.375
par-ilw	0.047	0.054	0.042	0.040	0.024	0.066
par-clw	0.060	0.091	0.193	0.242	0.393	0.234
Panel C: Portfolios calibrated nonparametrically						
Portfolios that consider the vector of means						
f-mv	0.128	0.205	0.647	1.153	8.432	3.019
Portfolios that do not consider the vector of means						
npar-lw	0.077	0.111	0.295	0.407	1.061	0.338
npar-ilw	0.038	0.035	0.032	0.034	0.023	0.062
npar-clw	0.061	0.084	0.174	0.215	0.345	0.218

Table B.2.: Turnover of shrinkage portfolios

This table reports the out-of-sample Turnover of the shrinkage portfolios. We consider an investor with a risk aversion parameter of $\gamma = 5$.

Policy	5IndP	10IndP	38IndP	48IndP	100FF	SP100
Panel A:Shrinkage portfolios with parametric calibration						
EQL Minimization						
mv-min	0.216	0.279	0.873	1.390	12.306	3.215
mv-ew	0.208	0.218	0.687	1.126	83.852	1.670
min-ew	0.086	0.124	0.335	0.463	3.207	1.344
Utility Maximization						
mv-min	0.229	0.295	0.801	1.214	7.260	2.796
mv-ew	0.186	0.245	0.723	1.084	40.024	1.464
min-ew	0.084	0.136	0.420	0.652	4.975	1.960
Variance Minimization						
mv-min	0.133	0.163	0.499	0.812	7.173	2.832
mv-ew	0.080	0.135	0.328	0.445	3.952	1.339
min-ew	0.084	0.136	0.420	0.652	4.974	1.960
Sharpe Ratio Maximization						
mv-min	0.210	0.262	0.945	1.729	145.685	3.155
mv-ew	0.208	0.290	1.002	1.704	12.673	2.412
min-ew	0.121	0.136	0.385	0.571	5.199	1.960
Panel B: Shrinkage portfolios with bootstrap calibration						
EQL Minimization						
mv-min	0.259	0.295	19.077	173.172	42.238	5.738
mv-ew	0.198	0.197	0.423	0.404	0.259	0.062
min-ew	0.083	0.111	0.199	0.162	0.023	0.062
Utility Maximization						
mv-min	0.259	0.312	35.319	14.670	29.513	3.097
mv-ew	0.169	0.216	0.420	0.342	0.108	0.062
min-ew	0.100	0.115	0.246	0.187	0.023	0.062
Variance Minimization						
mv-min	0.193	0.160	0.520	0.837	6.728	2.798
mv-ew	0.093	0.127	0.167	0.107	0.023	0.062
min-ew	0.098	0.116	0.244	0.189	0.023	0.062
Sharpe Ratio Maximization						
mv-min	0.282	0.273	1.180	2.702	281.946	4.940
mv-ew	0.205	0.287	0.791	0.835	0.024	0.063
min-ew	0.112	0.135	0.321	0.327	0.023	0.063
Panel C: Existing mixture of portfolios						
kz	0.210	0.262	0.945	1.729	149.144	3.155
tz	0.208	0.296	0.851	1.255	5.326	1.176
dm	0.147	0.151	0.386	0.571	5.280	1.960

B.3. Out-of-sample results for M=150

Table B.3 reports the annualized Sharpe ratio of returns, net of transaction costs of 50 basis points, of the benchmark portfolios and the portfolios constructed with the shrinkage moments. From Panel A we observe that the minimum-variance portfolio with the shrinkage covariance matrix proposed by Ledoit and Wolf (2004b) (lw) attains the highest out-of-sample Sharpe ratio among all benchmark portfolios. Panel B reports the Sharpe ratio for the portfolios formed with our studied shrinkage moments calibrated under the normality assumption. We observe that the minimum-variance portfolio formed from the shrinkage covariance matrix that accounts for the expected quadratic loss and the condition number (par-clw) outperforms the lw portfolio for medium and large datasets. This is because for medium and large datasets, the sample covariance matrix is more likely to be nearly singular. Also, we can observe that the differences between par-clw and lw are statistically significant for the 38IndP, 48IndP and 100FF datasets.

Table B.3.: Annualized Sharpe ratio of benchmark portfolios and portfolios with shrinkage moments (κ =50 basis points)

This table reports the out-of-sample annualized Sharpe ratio of benchmark portfolios and portfolios constructed by using the shrinkage moments. We adjust the Sharpe ratio with transaction costs, where we assume that the chargeable fee is equivalent to 50 basis points (bp). We consider an investor with a risk aversion parameter of $\gamma = 5$. One, two and three asterisks indicate that the difference with the lw portfolio is statistically different from zero for a 90%, 95% and 99% confidence interval, respectively.

Policy	5IndP	10IndP	38IndP	48IndP	100FF	SP100
Panel A: Benchmark Portfolios						
Portfolios that consider the vector of means						
mv	0.593**	0.599***	0.004***	-0.003***	-0.463***	-0.102*
bs	0.733*	0.817*	0.250***	0.164***	-0.343***	0.113
Portfolios that do not consider the vector of means						
min	0.841	0.945	0.528***	0.378***	-0.014***	0.399
lw	0.863	0.955	0.731	0.651	1.003	0.687
lw-m	0.836	0.953	0.649***	0.607	0.843***	0.648
Naïve Portfolios						
1/N	0.761	0.780	0.695	0.688	0.712	0.328
Panel B: Portfolios calibrated parametrically						
Portfolios that do not consider the vector of means						
par-lw	0.845	0.945	0.643***	0.553***	0.762***	0.641
par-ilw	0.877	0.907	0.716	0.693	0.713	0.337
par-clw	0.853	0.948	0.824***	0.794***	1.164***	0.700
Panel C: Portfolios calibrated nonparametrically						
Portfolios that consider the vector of means						
f-mv	0.809	0.881	0.403***	0.270***	0.018***	0.257
Portfolios that do not consider the vector of means						
npar-lw	0.860	0.954	0.711***	0.622***	0.929***	0.667
npar-ilw	0.848	0.844	0.701	0.690	0.712	0.328
npar-clw	0.863	0.954	0.858**	0.825***	1.152*	0.702

Panel C reports the portfolios constructed with our studied shrinkage moments calibrated without making any assumption about the distribution of stock returns. To estimate the shrinkage covariance matrix and the inverse covariance matrix, we apply the proposed smoothed bootstrap. First, we observe that the mean-variance portfolio obtained from our considered shrinkage vector of means beats the benchmark mean-variance portfolios (mv and bs) across every dataset. We also observe that the nonparametric calibration works better than the parametric approach to calibrate the shrinkage covariance matrix of the minimum-variance portfolios. In general, the nonparametric approach gives larger shrinkage intensities, which seems to imply that empirical data departs from the normality assumption, and therefore sample estimators require larger shrinkage intensities than those suggested by the parametric approach.

Table B.4 reports the annualized Sharpe ratio of returns net of transaction costs of 50 basis points of the shrinkage portfolios. Panel A reports the annualized adjusted Sharpe ratios of the shrinkage portfolios calibrated via parametric assumptions. Panel B reports the annualized adjusted Sharpe ratios of the shrinkage portfolios calibrated via bootstrap. Panel C reports the results of the shrinkage portfolios from the literature. From Panel A we make two observations. First, the variance minimization criterion is the best calibration criterion for the portfolios that consider the vector of means, mv-min and mv-ew, whereas the expected quadratic loss is the best calibration criterion for the portfolio that does not consider the vector of means, min-ew. Moreover, we observe that the best shrinkage portfolio is the min-ew portfolio. The explanation for this is that it is well-known that it is much harder to estimate the mean than the covariance matrix of asset returns from empirical data. Therefore, a mixture of the minimum-variance portfolio with the equally-weighted portfolio always outperforms any other combination that considers the vector of means, which would provide more unstable portfolios with lower adjusted Sharpe ratios.

From Panel B of Table B.4, we observe again that the best shrinkage portfolio is the mixture formed with the minimum-variance portfolio and the equally-weighted portfolio. Furthermore, we also observe that the expected quadratic loss minimization criterion is, in general, the best calibration criterion in terms of Sharpe ratio. The results obtained under the nonparametric bootstrap approach are, in general, slightly better than the results obtained under the assumption of normally distributed returns because empirical returns seem to depart from the normality assumption.

Panel C of Table B.4, shows the annualized Sharpe ratio of the existing mixture of portfolios from the literature. We observe that among the mixture of portfolios, the mixture formed by the minimum-variance portfolio and the equally weighted portfolio offers the best results. This mixture, however, performs worse than our studied shrinkage portfolio formed with the minimum-variance portfolio and the equally-weighted portfolio across every dataset. Thus, our proposed framework to construct shrinkage portfolios turns out to hedge better the investor's portfolio against estimation error.

Tables B.5 and B.6 report the results for the out-of-sample standard deviation of the studied portfolios. The results obtained from these tables are consistent with the results obtained for the Sharpe ratio and the turnover.

Summarizing, we observe that the main findings are in general the same as in the case of M=120, and in turn the qualitative results do not change for M=150. Our first observation is that portfolios computed from our studied shrinkage vector of means outperform those computed from the Bayes-Stein vector of means of Jorion (1986). Second, we observe that controlling for the condition number of the shrinkage covariance matrix results in portfolio weights that are more stable, and this leads to better adjusted Sharpe ratios for medium and large datasets. Third, for shrinkage portfolios that consider the vector of means, the variance minimization criterion is the most robust criterion, whereas for

Table B.4.: Annualized Sharpe ratio with transaction costs of shrinkage portfolios

This table reports the out-of-sample annualized Sharpe ratio (adjusted with 50 bp) of the shrinkage portfolios for an investor with $\gamma = 5$. One, two and three asterisks indicate that the difference with the lw portfolio is statistically different from zero for a 90%, 95% and 99% confidence interval, respectively.

Policy	5IndP	10IndP	38IndP	48IndP	100FF	SP100
Panel A: Shrinkage portfolio with parametric calibration						
EQL Minimization						
mv-min	0.702**	0.812**	0.313***	0.274**	-0.198***	0.307
mv-ew	0.647**	0.760*	0.483	0.437	0.028***	0.188*
min-ew	0.844	0.954	0.658	0.587	0.435***	0.509
Utility Maximization						
mv-min	0.708*	0.790*	0.303***	0.305**	-0.025***	0.388
mv-ew	0.691*	0.738*	0.386**	0.436	0.205***	0.234
min-ew	0.847	0.949	0.593***	0.472***	0.165***	0.483
Variance Minimization						
mv-min	0.820	0.925	0.527***	0.374***	-0.054***	0.414
mv-ew	0.773	0.796*	0.594	0.670	0.425***	0.277
min-ew	0.847	0.949	0.593***	0.472***	0.165***	0.483
Sharpe Ratio Maximization						
mv-min	0.714*	0.807**	0.284***	0.208**	-0.208***	0.273
mv-ew	0.685**	0.715**	0.249***	0.254**	-0.188***	0.111*
min-ew	0.817	0.955	0.606**	0.508***	0.104***	0.472
Panel B: Shrinkage portfolios with bootstrap calibration						
EQL Minimization						
mv-min	0.708**	0.795**	0.277***	-0.356***	-0.612***	0.398
mv-ew	0.657**	0.769*	0.575	0.596	0.713	0.329
min-ew	0.853	0.952	0.711	0.675	0.712	0.330
Utility Maximization						
mv-min	0.714*	0.771**	0.264***	0.158**	-0.471***	0.399
mv-ew	0.701**	0.750**	0.514	0.615	0.712*	0.328
min-ew	0.844	0.945	0.688	0.662	0.712	0.328
Variance Minimization						
mv-min	0.821	0.931	0.528***	0.374***	-0.093***	0.399
mv-ew	0.757	0.798	0.641	0.685	0.712	0.328
min-ew	0.855	0.948	0.684	0.660	0.712	0.328
Sharpe Ratio Maximization						
mv-min	0.702**	0.793**	0.261***	0.050***	-0.471***	0.153*
mv-ew	0.689**	0.714**	0.308***	0.401	0.714	0.331
min-ew	0.824	0.955	0.627	0.590	0.713	0.335
Panel C: Existing mixture of portfolios						
kz	0.714*	0.807**	0.284***	0.208**	-0.208***	0.273
tz	0.673**	0.704**	0.301***	0.360*	0.083***	0.195
dm	0.813	0.941	0.603**	0.508***	0.104***	0.472

shrinkage portfolios that do not consider the vector of means, the expected quadratic loss criterion works better. Finally, the studied nonparametric approach to calibrate shrinkage estimators captures the departure from normality in real return data and this results in more stable portfolios (small turnover) with reasonable Sharpe ratios.

Table B.5.: Standard deviation of benchmark portfolios and portfolios with shrinkage moments

This table reports the out-of-sample standard deviation of benchmark portfolios and portfolios constructed by using the shrinkage moments. We consider an investor with a risk aversion parameter of $\gamma = 5$. One, two and three asterisks indicate that the difference with the lw portfolio is statistically different from zero for a 90%, 95% and 99% confidence interval, respectively.

Policy	5IndP	10IndP	38IndP	48IndP	100FF	SP100
Panel A: Benchmark Portfolios						
Portfolios that consider the vector of means						
mv	0.161***	0.157***	0.244***	0.336***	0.417***	0.267***
bs	0.143**	0.134***	0.167***	0.224***	0.346***	0.200***
Portfolios that do not consider the vector of means						
min	0.138	0.126**	0.131***	0.137***	0.179***	0.171***
lw	0.136	0.124	0.120	0.124	0.125	0.122
lw-m	0.138	0.126	0.121	0.123	0.132***	0.120
Naïve Portfolios						
1/N	0.154***	0.148***	0.166***	0.165***	0.174***	0.169***
Panel B: Portfolios calibrated parametrically						
Portfolios that do not consider the vector of means						
par-lw	0.136	0.125	0.124***	0.128***	0.137***	0.125**
par-ilw	0.138**	0.132*	0.159***	0.161***	0.174***	0.167***
par-clw	0.136	0.124	0.119	0.122	0.122	0.121
Panel C: Portfolios calibrated nonparametrically						
Portfolios that consider the vector of means						
f-mv	0.138	0.128**	0.139***	0.169***	0.204***	0.178***
Portfolios that do not consider the vector of means						
npar-lw	0.136	0.124	0.121***	0.125***	0.128***	0.123*
npar-ilw	0.143***	0.140***	0.164***	0.164***	0.174***	0.169***
npar-clw	0.137	0.124	0.119	0.122	0.123	0.121

B.4. Out-of-sample results for M=60

In this part of the analysis, we report the out-of-sample results for estimated portfolio that shrink the covariance matrix. The estimation window is M=60 observations and because this is a very small sample to estimate portfolios with large number of asset, we do not report the results for the other considered portfolios in the chapter. In general, we again observe that the proposed calibration criterion that accounts for the condition number provides better results than all the benchmark portfolios: larger Sharpe ratio with moderate turnover and small volatility; see Tables B.7, B.8 and B.9. Also, we observe that the multivariate smoothed bootstrap tends to work better than the parametric approach, and this provides larger Sharpe ratios with lower turnover.

Table B.6.: Standard deviation of shrinkage portfolios

This table reports the out-of-sample standard deviation of the shrinkage portfolios. We consider an investor with a risk aversion parameter of $\gamma = 5$. One, two and three asterisks indicate that the difference with the lw portfolio is statistically different from zero for a 90%, 95% and 99% confidence interval, respectively.

Policy	5IndP	10IndP	38IndP	48IndP	100FF	SP100
Panel A: Shrinkage portfolios with parametric calibration						
EQL Minimization						
mv-min	0.145**	0.137***	0.159***	0.183***	0.251***	0.173***
mv-ew	0.152***	0.140***	0.157***	0.177***	0.216***	0.149*
min-ew	0.138	0.125	0.126**	0.130**	0.146***	0.134**
Utility Maximization						
mv-min	0.145***	0.137***	0.158***	0.177***	0.195***	0.170***
mv-ew	0.149***	0.140***	0.165***	0.186***	0.198***	0.147*
min-ew	0.137	0.125	0.127***	0.132***	0.160***	0.142***
Variance Minimization						
mv-min	0.141**	0.127*	0.131***	0.137***	0.184***	0.172***
mv-ew	0.146***	0.138***	0.155***	0.162***	0.174***	0.145
min-ew	0.137	0.125	0.127***	0.132***	0.160***	0.142***
Sharpe Ratio Maximization						
mv-min	0.144***	0.135***	0.161***	0.203***	0.277***	0.176***
mv-ew	0.150***	0.142***	0.181***	0.228***	0.297***	0.162***
min-ew	0.136	0.124	0.126**	0.132***	0.166***	0.142***
Panel B: Shrinkage portfolios with nonparametric calibration						
EQL Minimization						
mv-min	0.144***	0.137***	0.160***	0.335***	0.264***	0.171***
mv-ew	0.151***	0.140***	0.153***	0.160***	0.174***	0.168***
min-ew	0.137	0.125	0.132**	0.142***	0.174***	0.168***
Utility Maximization						
mv-min	0.144**	0.137***	0.161***	0.971***	0.183***	0.171***
mv-ew	0.149***	0.140***	0.155***	0.161***	0.174***	0.169***
min-ew	0.138***	0.125	0.129**	0.140***	0.174***	0.169***
Variance Minimization						
mv-min	0.140*	0.127**	0.131***	0.138***	0.181***	0.171***
mv-ew	0.148***	0.140***	0.159***	0.162***	0.174***	0.169***
min-ew	0.138**	0.125	0.129**	0.140***	0.174***	0.169***
Sharpe Ratio Maximization						
mv-min	0.145***	0.135***	0.166***	0.231***	0.373***	0.189***
mv-ew	0.149***	0.142***	0.171***	0.193***	0.174***	0.168***
min-ew	0.137	0.124	0.127**	0.135**	0.174***	0.167***
Panel C: Existing mixture of portfolios						
kz	0.144***	0.135***	0.161***	0.203***	0.277***	0.176***
tz	0.150***	0.143***	0.174***	0.202***	0.222***	0.149**
dm	0.136	0.125	0.126**	0.132***	0.166***	0.142***

Table B.7.: Sharpe ratio for estimation window M = 60 of benchmark portfolios and portfolios estimated with shrinkage moments

This table reports the out-of-sample Sharpe ratio net of transaction costs of benchmark portfolios and portfolios constructed by using the studied shrinkage estimators for the moments of asset returns.

Policy	5IndP	10IndP	38IndP	48IndP	100FF	SP100
Panel A: Benchmark Portfolios						
lw	0.798	0.884	0.538	0.403	0.647	0.862
lw-m	0.747	0.824	0.462	0.371	0.549	0.840
1/N	0.786	0.811	0.739	0.738	0.803	0.672
Panel B: Portfolios calibrated parametrically						
par-lw	0.789	0.866	0.450	0.265	0.527	0.798
par-clw	0.801	0.911	0.681	0.672	0.948	0.906
Panel C: Portfolios calibrated nonparametrically						
par-lw	0.793	0.877	0.503	0.338	0.578	0.813
par-clw	0.802	0.918	0.708	0.708	0.967	0.909

Table B.8.: Turnover of benchmark portfolios and portfolios estimated with shrinkage moments

This table reports the out-of-sample turnover of benchmark portfolios and portfolios constructed by using the studied shrinkage estimators for the moments of asset returns.

Policy	5IndP	10IndP	38IndP	48IndP	100FF	SP100
Panel A: Benchmark Portfolios						
lw	0.104	0.166	0.418	0.526	1.127	0.320
lw-m	0.173	0.243	0.474	0.513	1.240	0.303
1/N	0.017	0.024	0.030	0.032	0.023	0.059
Panel B: Portfolios calibrated parametrically						
par-lw	0.116	0.194	0.518	0.668	1.434	0.389
par-clw	0.092	0.140	0.277	0.297	0.474	0.249
Panel C: Portfolios calibrated nonparametrically						
par-lw	0.117	0.178	0.458	0.593	1.290	0.375
par-clw	0.093	0.130	0.254	0.277	0.353	0.198

B.5. Direct approach to shrink the covariance matrix using the condition number

In Section 2.2.1 we propose a criterion that takes explicitly into account the matrix condition number to calibrate the shrinkage covariance matrix. In particular, our proposed calibration criterion minimizes the expected quadratic loss and the condition number of the estimated covariance matrix. We solve the following optimization problem:

$$\alpha = \text{argmin}\left\{\delta_{\Sigma_{sh}} - \phi RIAL\left(\Sigma_{sh}\right)\right\}, \tag{B.2}$$

Table B.9.: Standard deviation of benchmark portfolios and portfolios estimated with shrinkage moments

This table reports the out-of-sample standard deviation of benchmark portfolios and portfolios constructed by using the studied shrinkage estimators for the moments of asset returns.

Policy	5IndP	10IndP	38IndP	48IndP	100FF	SP100
Panel A: Benchmark Portfolios						
lw	0.134	0.118	0.122	0.120	0.130	0.113
lw-m	0.133	0.121	0.122	0.117	0.131	0.112
1/N	0.153	0.146	0.166	0.166	0.174	0.151
Panel B: Portfolios calibrated parametrically						
par-lw	0.132	0.119	0.125	0.127	0.136	0.115
par-clw	0.133	0.118	0.120	0.117	0.124	0.113
Panel C: Portfolios calibrated nonparametrically						
par-lw	0.133	0.118	0.123	0.122	0.132	0.115
par-clw	0.135	0.118	0.120	0.117	0.128	0.114

where $\delta_{\Sigma_{sh}}$ represents the condition number of the shrinkage matrix and $RIAL\left(\Sigma_{sh}\right)$ is the relative improvement in average loss of the shrinkage covariance matrix.

We define parameter ϕ as the value that minimizes the out-of-sample portfolio variance, i.e. $\phi = \arg\min_{\phi} \sigma^2_{\phi}$, where σ^2_{ϕ} is the out-of-sample portfolio variance of the minimum variance portfolio formed with the shrinkage covariance matrix Σ_{sh}, calibrated by criterion (B.2). However, we can use a more direct approach where we set $\phi = 0$ ($\phi \to \infty$) when the out-of-sample portfolio variance of the equally-weighted portfolio is lower (larger) than the out-of-sample portfolio variance of the minimum-variance portfolio computed with the shrinkage covariance matrix that minimizes its expected quadratic loss.[1]

We compare the out-of-sample results of these two approaches in Table B.10. Portfolio par-clw is the minimum-variance portfolio that calibrates the shrinkage covariance matrix using (B.2), which we term the *mixed approach*, and portfolio da-clw is the *direct approach*. We observe that, in general, the mixed approach provides higher Sharpe ratios, net of transaction costs, and the out-of-sample standard deviation is lower than that of the portfolio calibrated by using the direct approach.

These results show that there is always a parameter $0 < \phi < \infty$ that provides better out-of-sample results than the direct approach for estimated portfolios constructed with a shrinkage covariance matrix.

[1] We implement the equally-weighted portfolio when $\phi = 0$ because in that case the resulting covariance matrix from (B.2) provides the equally-weighted portfolio. On the other hand, we implement the minimum-variance portfolio constructed with the shrinkage covariance matrix that minimizes the expected quadratic loss when $\phi \to \infty$ because in that situation, the resulting shrinkage covariance matrix from (B.2) is the shrinkage covariance matrix that minimizes the expected quadratic loss.

Table B.10.: Mixed approach vs Direct approach

This table reports the out-of-sample annualized net Sharpe ratio and the standard deviation of minimum-variance portfolios computed with the shrinkage covariance matrix that accounts for the condition number. We compare the results between the mixed approach and the direct approach. We assume that transaction costs are equal to 50 basis points (bp).

Policy	5IndP	10IndP	38IndP	48IndP	100FF	SP100
Panel A: Net Sharpe ratio						
par-clw	0.890	0.956	0.823	0.792	1.194	0.622
da-clw	0.891	0.944	0.645	0.559	0.687	0.563
Panel B: Standard deviation						
par-lw	0.137	0.124	0.122	0.125	0.121	0.126
da-clw	0.137	0.125	0.126	0.134	0.142	0.132

Bibliography

[1] Balduzzi, P. and A. W. Lynch (1999). Transaction costs and predictability: some utility cost calculations. *Journal of Financial Economics 52*(1), 47–78.

[2] Barry, C. B. (1974). Portfolio analysis under uncertain means, variances, and covariances. *The Journal of Finance 29*(2), 515–522.

[3] Basak, G. K., R. Jagannathan, and T. Ma (2009). Jackknife estimator for tracking error variance of optimal portfolios. *Management Science 55*(6), 990–1002.

[4] Bawa, V. S., S. J. Brown, and R. W. Klein (1979). *Estimation Risk and Optimal Portfolio Choice*. North-Holland Pub. Co. (Amsterdam and New York and New York).

[5] Best, M. J. and R. R. Grauer (1991). On the sensitivity of mean-variance-efficient portfolios to changes in asset means: Some analytical and computational results. *The Review of Financial Studies 4*, 315–342.

[6] Best, M. J. and R. R. Grauer (1992). Positively weighted minimum-variance portfolios and the structure of asset expected returns. *The Journal of Financial and Quantitative Analysis 27*(4), 513–537.

[7] Britten-Jones, M. (1999). The sampling error in estimates of mean-variance efficient portfolio weights. *Journal of Finance 54*(2), 655–671.

[8] Broadie, M. (1993). Computing efficient frontiers using estimated parameters. *Annals of Operations Research 45*, 21–58.

[9] Chan, L. C., J. Karceski, and J. Lakonishok (1999). On portfolio optimization: Forecasting covariances and choosing the risk model. *Review of Financial Studies 12*, 937–74.

[10] Constantinides, G. M. (1979). Multiperiod consumption and investment behavior with convex transactions costs. *Management Science 25*(11), pp. 1127–1137.

[11] Cornuejols, G. and R. Tutuncu (2007). *Optimization Methods in Finance*. Cambridge University Press.

[12] Davis, M. H. A. and A. R. Norman (1990). Portfolio selection with transaction costs. *Mathematics of Operations Research 15*(4), pp. 676–713.

[13] DeMiguel, V., L. Garlappi, F. J. Nogales, and R. Uppal (2009). A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms. *Management Science 55*, 798–812.

[14] DeMiguel, V., L. Garlappi, and R. Uppal (2009). Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *Review of Financial Studies 22*(5), 1915–1953.

[15] DeMiguel, V., A. Martin-Utrera, and F. J. Nogales (2013). Size matters: Optimal calibration of shrinkage estimators for portfolio selection. *Journal of Banking and Finance 37*(8), 3018–3034.

[16] DeMiguel, V. and F. J. Nogales (2009). Portfolio selection with robust estimation. *Operations Research 57*, 560–577.

[17] Dey, D. K. and C. Srinivasan (1985). Estimation of a covariance matrix under Stein's loss. *The Annals of Statistics 13*(4), 1581–1591.

[18] Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics 7*(1), 1–26.

[19] Efron, B. and G. Gong (1983). A leisurely look at the bootstrap, the jackknife, and cross-validation. *The American Statistician 37*(1), 36–48.

[20] Efron, B. and R. Tibshirani (1993). *An Introduction to the Bootstrap*. Chapman & Hall.

[21] Engle, R. and R. Ferstenberg (2007). Execution risk. *The Journal of Portfolio Management 33*, 34–44.

[22] Engle, R., R. Ferstenberg, and J. Russell (2012). Measuring and modeling execution cost and risk. *The Journal of Portfolio Management 38*, 14–28.

[23] Frahm, G. and C. Memmel (2010). Dominating estimators for minimum-variance portfolios. *Journal of Econometrics 159*(2), 289–302.

[24] Frost, P. A. and J. E. Savarino (1986). An empirical bayes approach to efficient portfolio selection. *The Journal of Financial and Quantitative Analysis 21*(3), 293–305.

[25] Garlappi, L., R. Uppal, and T. Wang (2007). Portfolio selection with parameter and model uncertainty: A multi-prior approach. *Review of Financial Studies 20*, 41–81.

[26] Garleanu, N. and L. H. Pedersen (2012). Dynamic trading with predictable returns and transaction costs. *The Journal of Finance*, Forthcoming.

[27] Goldfarb, D. and G. Iyengar (2003). Robust portfolio selection problems. *Mathematics of Operations Research 28*(1), 1–38.

[28] Greene, W. H. (2003). *Econometrics Analysis, Fifth Edition*. Prentice Hall.

[29] Greenwood, R. (2005). Short- and long-term demand curves for stocks: Theory and evidence on the dynamics of arbitrage. *Journal of Financial Economics 75*, 607–649.

[30] Haff, L. R. (1979). An identity for the wishart distribution with applications. *Journal of Multivariate Analysis 9*, 531–544.

[31] Jagannathan, R. and T. Ma (2003). Risk reduction in large portfolios: Why imposing the wrong constraints helps. *The Journal of Finance 58*, 1651–1684.

[32] James, W. and J. Stein (1961). Estimation with quadratic loss. *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, 361–379.

[33] Jobson, J. D. and B. Korkie (1981). Putting Markowitz theory to work. *Journal of Portfolio Management 7*, 70–74.

[34] Jorion, P. (1986). Bayes-Stein estimation for portfolio analysis. *The Journal of Financial and Quantitative Analysis 21*(3), 279–292.

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

[36] Kourtis, A., G. Dotsis, and R. N. Markellos (2012). Parameter uncertainty in portfolio selection: Shrinking the inverse covariance matrix. *Journal of Banking and Finance 36*(9), 2522–2531.

[37] Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica 53*(6), 1315–1335.

[38] Ledoit, O. and M. Wolf (2003). Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of Empirical Finance 10*, 603–621.

[39] Ledoit, O. and M. Wolf (2004a). Honey, I shrunk the sample covariance matrix. *Journal of Portfolio Management 30*, 110–119.

[40] Ledoit, O. and M. Wolf (2004b). A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis 88*, 365–411.

[41] Ledoit, O. and M. Wolf (2008). Robust performance hypothesis testing with the sharpe ratio. *Journal of Empirical Finance 15*, 850–859.

[42] Ledoit, O. and M. Wolf (2011). Robust performance hypothesis testing with the variance. *Wilmott Magazine forthcoming.*

[43] Lillo, F., J. D. Farmer, and R. N. Mantegna (2003). Master curve for price-impact function. *Nature 421*, 129–130.

[44] Liu, H. (2004). Optimal consumption and investment with transaction costs and multiple risky assets. *Journal of Finance 59*(1), 289–338.

[45] Liu, H. and M. Loewenstein (2002). Optimal portfolio selection with transaction costs and finite horizons. *Review of Financial Studies 15*(3), 805–835.

[46] MacKinlay, A. C. and L. Pastor (2000). Asset pricing models: Implications for expected returns and portfolio selection. *Review of Financial studies 13*, 883–916.

[47] Markowitz, H. (1952). Portfolio selection. *The Journal of Finance 7*(1), 77–91.

[48] Merton, R. C. (1980). On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics 8*, 323–361.

[49] Michaud, R. C. (1998). *Efficient Asset Management.* Oxford University Press.

[50] Michaud, R. O. (1989). The markowitz optimization enigma: Is 'optimized' optimal? *Financial Analysts Journal 45*(1), 31–42.

[51] Pastor, L. (2000). Portfolio selection and asset pricing models. *The Journal of Finance 55*, 179–223.

[52] Pastor, L. and R. F. Stambaugh (2000). Comparing asset pricing models: An investment perspective. *Journal of Financial Economics 56*, 335–381.

[53] Politis, D. N. and J. P. Romano (1994). The stationary bootstrap. *Journal of the American Statistical Association 89*(428), pp. 1303–1313.

[54] Rustem, B., R. G. Becker, and W. Marty (2000). Robust min-max portfolio strategies for rival forecast and risk scenarios. *Journal of Economic Dynamics and Control 24*, 1591–1621.

[55] Stein, C. (1975). Estimation of a covariance matrix. *Rietz Lecture, 39th Annual Meeting IMS, Atlanta, GA.*

[56] Tu, J. and G. Zhou (2011). Markowitz meets Talmud: A combination of sophisticated and naive diversification strategies. *Journal of Financial Economics 99*(1), 204 – 215.

[57] Tutuncu, R. H. and M. Koenig (2004). Robust asset allocation. *Annals of Operations Resarch 132*, 157–187.

[58] Wang, Z. (2005). A shrinkage approach to model uncertainty and asset allocation. *Review of Financial Studies 18*(2), 673–705.

[59] Zumbach, G. (2009). Inference on multivariate arch processes with large sizes. *SSRN, eLibrary*.