

Parameter Uncertainty in Multiperiod Portfolio Optimization with Transaction Costs

Victor DeMiguel, Alberto Martín-Utrera, and Francisco J. Nogales*

Abstract

We study the impact of parameter uncertainty on the expected utility of a multiperiod investor subject to quadratic transaction costs. We characterize the utility loss associated with ignoring parameter uncertainty, and show that it is equal to the product between the single-period utility loss and another term that captures the effects of the multiperiod mean-variance utility and transaction cost losses. To mitigate the impact of parameter uncertainty, we propose two multiperiod shrinkage portfolios and demonstrate with simulated and empirical data sets that they substantially outperform portfolios that ignore parameter uncertainty, transaction costs, or both.

I. Introduction

Markowitz (1952) shows that an investor who cares only about the portfolio mean and variance should hold a portfolio on the efficient frontier. Markowitz's mean-variance framework is the cornerstone of most practical investment approaches, but it relies on three restrictive assumptions. First, the investor is *myopic* and maximizes a single-period utility. Second, financial markets are frictionless. Third, the investor knows the exact parameters that capture asset price dynamics. However, these assumptions are unrealistic.

There is 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

*DeMiguel, avmiguel@london.edu, London Business School, London NW1 4SA, United Kingdom; Martín-Utrera (corresponding author), a.martinutrera@lancaster.ac.uk, Lancaster University Management School, Lancaster LA1 4YX, United Kingdom; and Nogales, fcojavier.nogales@uc3m.es, Universidad Carlos III de Madrid, Getafe (Madrid) 28903, Spain. We thank an anonymous referee, Stephen Brown (the editor), Raymond Kan, Raman Uppal, Grigory Vilkov, Guofu Zhou, and seminar participants at Universidad Carlos III de Madrid, Birkbeck College, University of Bristol, Essex Business School, Surrey Business School, Manchester Business School, University of Tennessee, and the 2012 Institute for Operations Research and the Management Sciences Annual Meeting. Martín-Utrera and Nogales gratefully acknowledge financial support from the Spanish government through project MTM2010-16519.

is characterized by a no-trade interval. The case with multiple risky assets and proportional transaction costs is generally intractable analytically.¹ Garleanu and Pedersen (GP) (2013) 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.

There is also extensive literature on parameter uncertainty in portfolio selection for the case of a single-period investor who *is not subject to transaction costs*. This literature includes Bayesian approaches with diffuse priors (Klein and Bawa (1976), Brown (1978)), Bayesian approaches with priors based on asset pricing models (MacKinlay and Pástor (2000), Pástor (2000), and Pástor and Stambaugh (2000)), Bayesian approaches with priors based on economic objectives (Tu and Zhou (2010)), shrinkage approaches (Ledoit and Wolf (2004)), robust optimization methods (Cornuejols and Tutuncu (2007), Goldfarb and Iyengar (2003), Garlappi, Uppal, and Wang (2007), Rustem, Becker, and Marty (2000), and Tutuncu and Koeing (2004)), Bayesian robust optimization (Wang (2005)), mean-variance timing rules (Kirby and Ostdiek (2012)), and methods based on imposing constraints (Best and Grauer (1992), Jagannathan and Ma (2003), and DeMiguel, Garlappi, Nogales, and Uppal (2009)).

Kan and Zhou (2007) characterize analytically the utility loss of a mean-variance investor who suffers from parameter uncertainty. Moreover, they propose two single-period shrinkage portfolios that shrink the sample mean-variance portfolio toward a target portfolio, and they analytically characterize the shrinkage intensities that minimize the investor's utility loss from parameter uncertainty.²

In this paper, we consider the impact of parameter uncertainty on the performance of a multiperiod mean-variance investor facing quadratic transaction costs. Our first contribution is to give a closed-form expression for the utility loss of an investor who uses sample information to construct his or 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 coming from the multiperiod mean-variance utility and the multiperiod transaction costs.

We also use our characterization of the utility loss to understand how 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. The explanation for this is that transaction costs induce the investor to trade at a slower rate, and as a result the impact of estimation error is postponed to future time periods that have a smaller impact on the overall discounted utility. We also find that an investor with a 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

¹Liu (2004), however, characterizes analytically the case where asset returns are uncorrelated and the investor has constant absolute risk aversion (CARA) utility.

²See also Tu and Zhou (2010), who consider a combination of the sample mean-variance portfolio with the equal-weighted portfolio.

impatient, the cost of making a trade is relatively more important as compared with the future expected payoff of the corresponding trade. Hence, larger trading costs or a 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 two shrinkage approaches designed to combat estimation risk in the multiperiod mean-variance framework with quadratic transaction costs. From GP, it is easy to show that, in the absence of estimation error, the optimal portfolio policy is to trade toward the Markowitz portfolio at a fixed trading rate every period. Our first shrinkage approach consists of shrinking the Markowitz portfolio toward a target that is less sensitive to estimation error, while maintaining the trading rate fixed at its nominal value. This portfolio aims to diversify the effects of estimation risk across different funds. The second approach consists of shrinking the trading rate, in addition to shrinking the Markowitz portfolio. This portfolio aims to smooth the investor trading activity to avoid extreme positions that may result in extreme negative outcomes due to the effects of parameter uncertainty.

We consider two variants of the first approach. The first variant consists of shrinking the Markowitz portfolio toward the risk-free asset. We term this portfolio a multiperiod 3-fund shrinkage portfolio because it is a combination of the investor's initial portfolio, the Markowitz portfolio, and the risk-free asset. The second variant shrinks the Markowitz portfolio toward the minimum-variance portfolio, and we term the resulting trading strategy as multiperiod 4-fund portfolio because it is a combination of the investor's initial portfolio, the Markowitz portfolio, the minimum-variance portfolio, and the risk-free asset. We show that the optimal shrinkage intensities for the 3- and 4-fund multiperiod 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 shrinkage approach, the nominal trading rate given by GP may not be optimal in the presence of parameter uncertainty. Hence, we propose versions of the 4-fund portfolio in which 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 identify the conditions in which the investor can obtain gains by shrinking the trading rate.

For tractability, the analysis just described relies on the assumption that the investor uses a fixed estimation window to construct his or her lifetime portfolio policy. Our third contribution is to relax this assumption by considering the case in which the investor uses expanding estimation windows; that is, the investor uses all available data at each point in time for estimation purposes. We find that for the case with expanding windows, the utility loss is no longer separable into the product of the single-period utility loss and a multiperiod factor. Nevertheless, we conjecture an approximation to the investor's expected loss that allows us to compute the optimal shrinkage intensities of the multiperiod shrinkage portfolios, and show with simulated data that the conjectured approximation is very accurate.

Our fourth contribution is to evaluate the out-of-sample performance of the proposed multiperiod shrinkage portfolios on simulated and empirical data sets. We find that the 4-fund portfolios (either with nominal or shrunk 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 under certain circumstances. Finally, we identify the situations when using expanding windows helps to improve performance.

The outline of the paper is as follows: Section II characterizes the expected loss of an investor who uses sample information to estimate his or her optimal trading strategy. Section III introduces the multiperiod shrinkage portfolios that help to reduce the effects of estimation risk, and Section IV studies the case with expanding windows. Section V tests the out-of-sample performance of our proposed multiperiod portfolios on simulated and empirical data sets. Section VI concludes.

II. Multiperiod Utility Loss

We adopt the framework proposed by GP. In this framework, the investor maximizes his or her multiperiod mean-variance utility, net of quadratic transaction costs, by choosing the number of shares to hold of each of the N risky assets. We focus on the case in which price changes in excess of the risk-free asset are independent and identically distributed (IID) as a normal distribution with mean μ and covariance matrix Σ .³

The investor's objective is

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

where $x_i \in \mathbb{R}^N$ for $i \geq 0$ contains the number of shares held of each of the N risky assets at time i , ρ is the investor's impatience factor, and γ is the absolute risk-aversion parameter. The term $(\lambda/2) \Delta x_i' \Sigma \Delta x_i$ is the quadratic transaction cost at the i th 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 i th 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 GP's assumption that the quadratic transaction costs are proportional to the covariance matrix Σ . GP provide microfoundations to justify the use of quadratic transaction costs, and we

³GP consider the case of predictable price changes; however, we focus on the IID multivariate normal case, which is customary in the transaction costs literature (see Constantinides (1979), Davis and Norman (1990), and Liu (2004)).

⁴Several authors have shown that the quadratic form matches the market impact costs observed in empirical data (see, e.g., Lillo, Farmer, and Mantegna (2003), Engle, Ferstenberg, and Russell (2012)).

can also find more examples in the literature that address this type of cost. For instance, 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.

GP show that the optimal multiperiod portfolio is a convex combination between the investor's current portfolio and the static mean-variance (Markowitz) portfolio,

$$(2) \quad x_i = (1 - \beta)x_{i-1} + \beta x^M,$$

where $\beta = \left(\sqrt{(\gamma + \tilde{\lambda}\rho)^2 + 4\gamma\lambda} - (\gamma + \tilde{\lambda}\rho) \right) / (2\lambda) \leq 1$ represents the investor's trading rate, $x^M = (1/\gamma)\Sigma^{-1}\mu$ is the Markowitz portfolio, and $\tilde{\lambda} = \lambda/(1 - \rho)$. Hence, it is optimal to invest in the static mean-variance portfolio, but it is prohibitive to trade toward the Markowitz portfolio in a single period, and thus the investor converges smoothly at a constant trading rate. The trading rate β increases with the absolute risk-aversion parameter γ , and decreases with the transaction cost parameter λ and the investor's impatience factor ρ (see GP).⁵

In a real-world application, investors ignore the true inputs that define the model. Therefore, it is interesting to characterize the expected loss for an investor who uses historical data to construct the optimal trading strategy (i.e., the *plug-in* approach). Specifically, let r_t for $t = 1, 2, \dots, T$ be the sample of excess price changes. Then, we consider the following sample estimators of the mean and covariance matrix:

$$(3) \quad \hat{\mu} = \frac{1}{T} \sum_{t=1}^T r_t, \quad \text{and} \quad \hat{\Sigma} = \frac{1}{T - N - 2} \sum_{t=1}^T (r_t - \hat{\mu})^2.$$

Kan and Zhou (2007) characterize the investor's expected utility loss as the difference between the single-period utility evaluated for the true Markowitz portfolio x^M and the expected single-period utility evaluated for the estimated Markowitz portfolio \hat{x}^M . They show that this is defined by $L_1(x^M, \hat{x}^M) = (1/2\gamma)[(c - 1)\theta + c(N/T)]$, where $c = [(T - N - 2)(T - 2)]/[(T - N - 1)(T - N - 4)]$, and $\theta = \mu'\Sigma^{-1}\mu$ is the squared Sharpe ratio for a static mean-variance investor.⁶

Similar to Kan and Zhou (2007), we study the investor's multiperiod expected loss as the difference between the investor's utility evaluated for the true optimal trading strategy x_i and the investor's expected utility evaluated for the estimated optimal trading strategy \hat{x}_i . We provide this result in the following proposition (*all proofs are given in the Appendix*).

⁵GP prove the monotonicity properties of β only for γ and λ . However, it is straightforward to prove the monotonicity of β with respect to ρ using their same arguments.

⁶The single-period utility loss is stated here in terms of the unbiased estimator of the Markowitz portfolio, whereas Kan and Zhou (2007) cast it in terms of the standard estimator.

Proposition 1. The expected loss of a multiperiod mean-variance investor is equal to the product between the utility loss of a single-period investor $L_1(x^M, \hat{x}^M)$ and a multiperiod term,

$$(4) \quad L(\{x_i\}, \{\hat{x}_i\}) = L_1(x^M, \hat{x}^M) \times [f_{mv} + f_{tc}],$$

where f_{mv} is the multiperiod mean-variance loss factor, and f_{tc} is the multiperiod transaction cost loss factor:

$$(5) \quad f_{mv} = \frac{1-\rho}{\rho} + \frac{(1-\rho)(1-\beta)^2}{1-(1-\rho)(1-\beta)^2} - 2 \frac{(1-\rho)(1-\beta)}{1-(1-\rho)(1-\beta)},$$

$$(6) \quad f_{tc} = \frac{\lambda}{\gamma} \frac{\beta^2}{1-(1-\rho)(1-\beta)^2}.$$

To understand why we name f_{mv} and f_{tc} as the multiperiod *mean-variance* and *transaction cost* factors, respectively, note that from the proof of Proposition 1 it is easy to see that the multiperiod expected loss can be written as $L = L_{mv} + L_{tc}$, where

$$(7) \quad L_{mv} \equiv L_1(x^M, \hat{x}^M) \times f_{mv} = \frac{\gamma}{2} \sum_{i=0}^{\infty} (1-\rho)^{i+1} E[\hat{x}'_i \Sigma \hat{x}_i - x'_i \Sigma x_i],$$

$$(8) \quad L_{tc} \equiv L_1(x^M, \hat{x}^M) \times f_{tc} \\ = \frac{\tilde{\lambda}}{2} \sum_{i=0}^{\infty} (1-\rho)^{i+1} E[\Delta \hat{x}'_i \Sigma \Delta \hat{x}_i - \Delta x'_i \Sigma \Delta x_i].$$

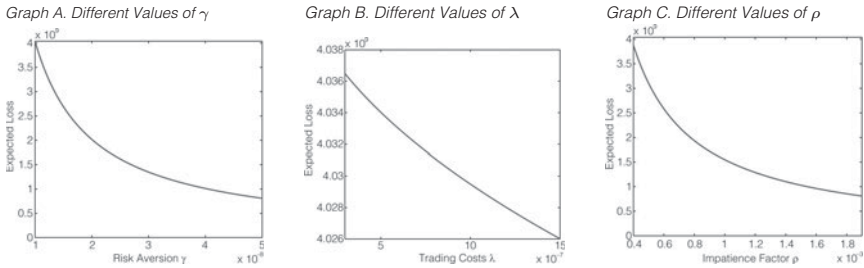
Moreover, equation (7) shows that the term L_{mv} depends only on the multiperiod *mean-variance* loss of the plug-in multiperiod portfolio \hat{x}_i , and equation (8) shows that the term L_{tc} depends only on the multiperiod *transaction cost* loss of the plug-in multiperiod portfolio. Therefore, we can say that the multiperiod mean-variance loss factor f_{mv} captures the multiperiod losses originating from the mean-variance utility, and the multiperiod transaction cost loss factor f_{tc} captures the multiperiod losses originating from the transaction costs.

For tractability, in Proposition 1, we assume the investor uses a *fixed* estimation window (from time $t = 1$ to time $t = T$) to construct his or her lifetime optimal portfolio policy. In Section IV, however, we relax this assumption by considering an investor who updates his or her portfolio policy at every time period to take into account every available observation; that is, we consider an investor who uses *expanding* estimation windows. Moreover, for the out-of-sample evaluation in Section V, we consider the cases with a rolling estimation window (i.e., an estimation window that considers at each time period the last T observations) and an expanding window (i.e., a window that considers all available observations from $t = 1$).

Finally, we study how the multiperiod expected loss depends on the absolute risk-aversion parameter γ , the transaction cost parameter λ , and the discount factor ρ . Figure 1 shows the results for the case in which the investor constructs the optimal trading strategy with $T = 500$ observations, and in which the population parameters μ and Σ are equal to the sample moments of the empirical data

FIGURE 1
Utility Loss of Multiperiod Investor

Figure 1 depicts the investor's expected utility loss for different values of the absolute risk-aversion parameter γ , the transaction cost parameter λ , and the impatience factor ρ . Our base-case investor has $\gamma = 10^{-8}$, $\lambda = 3 \times 10^{-7}$, and $\rho = 1 - \exp(-0.1/260)$. We consider an investor who uses 500 observations to construct the optimal trading strategy. The population parameters are defined with the sample moments of the empirical data set formed with commodity futures that we consider in Section V.



set of commodity futures described in Section V. We obtain three main insights from Figure 1. First, the multiperiod expected loss decreases with the absolute risk-aversion parameter γ . As in the static case, this is an intuitive result because as the investor becomes more risk averse, the investor's exposure to risky assets is lower, and thus the impact of parameter uncertainty is also smaller. Second, the multiperiod expected loss decreases with transaction costs λ . As trading costs increase, the optimal trading rate decreases, and thus the investor optimally slows the convergence to the Markowitz portfolio. This effect results in a delay of the impact of parameter uncertainty to future stages in which the overall importance of utility losses is smaller. Third, the multiperiod expected loss decreases with ρ . Roughly speaking, the investor's impatience factor ρ has a similar effect on the investor's expected utility to that of trading costs λ . When the investor is more impatient, the cost of making a trade has greater importance than the future expected payoff of the corresponding trade. Accordingly, a more impatient investor also postpones the impact of parameter uncertainty to future stages that have a lower impact on the overall investor's utility, which results in a lower expected loss.

We now study a single-period example for which we can analytically characterize the monotonicity properties of the expected utility loss with respect to γ , λ , and ρ .

Example 1. Consider a single-period mean-variance investor subject to quadratic transaction costs and whose initial portfolio is x_{-1} . The investor's decision problem is:

$$(9) \quad \max_x (1 - \rho) \left\{ x' \mu - \frac{\gamma}{2} x' \Sigma x \right\} - \frac{\lambda}{2} \Delta x' \Sigma \Delta x.$$

Notice that expression (9) is a good approximation to equation (1) when ρ is close to 1. From the first-order optimality conditions, it is easy to see that the investor's optimal portfolio is $x = (1 - \beta_1)x_{-1} + \beta_1 x^M$, where $\beta_1 = \gamma / (\gamma + \tilde{\lambda})$ is the

single-period trading rate. Substituting μ and Σ with their sample counterparts, it is easy to show that the investor's expected utility loss is

$$(10) \quad \beta_1 \times L_1(x^M, \hat{x}^M).$$

Moreover, it is straightforward to show that the single-period trading rate β_1 satisfies the same monotonicity properties as the multiperiod trading rate. As a result, the investor's expected loss is monotonically decreasing in the transaction cost parameter and the investor's impatience factor because β_1 is decreasing in these parameters and $L_1(x^M, \hat{x}^M)$ does not depend on them. On the other hand, we observe that as γ increases, β_1 increases and $L_1(x^M, \hat{x}^M)$ decreases. However, the overall impact of γ on the investor's expected loss is determined by $\beta_1 \times (1/\gamma) = 1/(\gamma + \tilde{\lambda})$, which is a decreasing function of γ . Thus, the investor's expected loss also decreases with the investor's absolute risk-aversion parameter γ .

III. Multiperiod Shrinkage Portfolios

In this section we propose two shrinkage approaches to mitigate the impact of estimation error on the multiperiod mean-variance utility of an investor who faces quadratic transaction costs. For tractability, in this section we assume that the investor uses a fixed estimation window, but in Section IV we show how to relax this assumption and consider the case with expanding windows. Section III.A discusses the first approach, which consists of shrinking the estimated Markowitz portfolio toward a target that is less sensitive to estimation error, while maintaining the trading rate fixed to its nominal value. Section III.B discusses the second approach, which, in addition, shrinks the trading rate.

A. Shrinking the Markowitz Portfolio

The optimal portfolio at period i , in the absence of estimation error, can be written as

$$(11) \quad x_i = (1 - \beta)^{i+1} x_{-1} + \beta x^M \sum_{j=0}^i (1 - \beta)^j.$$

Therefore, the true optimal multiperiod trading strategy allocates the investor's wealth into three funds: the risk-free asset, the initial portfolio x_{-1} , and the Markowitz portfolio. In the presence of parameter uncertainty the investor suffers from estimation error, which results in utility losses. A simple rule to minimize utility losses is to shrink the sample Markowitz portfolio toward a target portfolio that is less sensitive to estimation error. For the single-period case, Kan and Zhou (2007) show that this helps to mitigate the impact of parameter uncertainty.

We generalize the analysis of Kan and Zhou (2007) to the multiperiod case. In particular, we consider two novel multiperiod portfolios that maximize the investor's expected utility by shrinking the Markowitz portfolio toward a target portfolio. The first portfolio shrinks the Markowitz portfolio toward the portfolio that invests solely in the risk-free asset (i.e., toward a portfolio with $x = 0$ holdings).

Shrinking the Markowitz portfolio gives a portfolio in the ex ante sample capital market line, and the resulting trading strategy is as follows:

$$(12) \quad \hat{x}_i^{3F} = (1 - \beta)\hat{x}_{i-1}^{3F} + \beta\eta\hat{x}^M,$$

where η is the shrinkage intensity. We term this portfolio as the multiperiod 3-fund shrinkage portfolio because it invests in the risk-free asset, the investor's initial portfolio, and the sample Markowitz portfolio.

Second, we consider a multiperiod portfolio that combines the sample Markowitz portfolio with the sample minimum-variance portfolio $\hat{x}^{\text{Min}} = (1/\gamma)\hat{\Sigma}^{-1}\iota$, which is known to be less sensitive to estimation error than the mean-variance portfolio (see Kan and Zhou (2007)):⁷

$$(13) \quad \hat{x}_i^{4F} = (1 - \beta)\hat{x}_{i-1}^{4F} + \beta(\varsigma_1\hat{x}^M + \varsigma_2\hat{x}^{\text{Min}}),$$

where ς_1 and ς_2 are the shrinkage intensities for the Markowitz portfolio and the minimum-variance portfolio, respectively. We term the resulting trading strategy as the multiperiod 4-fund portfolio because it invests in the risk-free asset, the investor's initial portfolio, the sample Markowitz portfolio, and the sample minimum-variance portfolio.

Note that whereas 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 for the single-period case, but the following proposition shows that the optimal shrinkage intensities for the single-period and multiperiod cases coincide.

Proposition 2. The optimal shrinkage intensities for the 3- and 4-fund portfolios that minimize the utility loss of a multiperiod mean-variance investor $L(\{x_i\}, \{\hat{x}_i\})$ coincide with the optimal shrinkage intensities for the single-period investor who ignores transaction costs. Specifically, the optimal shrinkage intensity for the 3-fund portfolio η and the optimal shrinkage intensities for the 4-fund portfolio ς_1 and ς_2 are

$$(14) \quad \eta = c^{-1} \frac{\theta}{\theta + \frac{N}{T}},$$

$$(15) \quad \varsigma_1 = c^{-1} \frac{\Psi^2}{\Psi^2 + \frac{N}{T}},$$

$$(16) \quad \varsigma_2 = c^{-1} \frac{\frac{N}{T}}{\Psi^2 + \frac{N}{T}} \times \frac{\mu' \Sigma^{-1} \iota}{\iota' \Sigma^{-1} \iota},$$

where $\Psi^2 = \mu' \Sigma^{-1} \mu - (\mu' \Sigma^{-1} \iota)^2 / (\iota' \Sigma^{-1} \iota) > 0$.

⁷Notice that the minimum-variance portfolio does not depend on γ . However, for notational convenience, we multiply the unscaled minimum-variance portfolio with $(1/\gamma)$.

From Proposition 2, we observe that the optimal combination parameters ς_1 and ς_2 are independent of trading costs. The intuition for this result can be traced back to the optimal multiperiod trading strategy, which is a convex combination between the investor's current portfolio and the static mean-variance portfolio. In the presence of parameter uncertainty, the investor should choose a combination of the current portfolio and the portfolio that is optimal in the single-period framework with the parameter uncertainty of Kan and Zhou (2007).

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 3-fund shrinkage portfolio is also inadmissible in the sense that it is always optimal to shrink the Markowitz portfolio toward 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 (i.e., $\eta < 1$). Moreover, it is always optimal to combine the Markowitz portfolio with the target minimum-variance portfolio (i.e., $\varsigma_2 > 0$).

We use the commodity futures data set described in Section V to illustrate the benefits of the use of the multiperiod 3- and 4-fund shrinkage portfolios. We consider the base-case investor described in Section V, and assume that he or she constructs the optimal trading strategy with $T = 500$ observations. Moreover, we set the population parameters μ and Σ equal to the sample moments of the empirical data set of commodity futures. We study the investor's loss relative to the *true* investor's utility, and we find that the *relative loss* of the base-case investor who is using the 3-fund shrinkage portfolio in equation (12) is about *eight times smaller* than that of the base-case investor using the plug-in multiperiod portfolio. Also, the *relative loss* of the 4-fund shrinkage portfolio in equation (13) is about *11% smaller* than that of the 3-fund portfolio in equation (12). These results confirm that there is a clear advantage to the use of the 4-fund shrinkage portfolio with respect to the plug-in multiperiod portfolio and the multiperiod 3-fund shrinkage portfolio.

B. Shrinking the Trading Rate

In this section we study the additional utility gain associated with shrinking the trading rate in addition to the Markowitz portfolio. For the proposed shrinkage portfolios in equations (12) and (13), note that the *nominal* trading rate β in equation (2) may not be optimal in the presence of parameter uncertainty. We now consider optimizing 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 4-fund shrinkage portfolio in equation (13) may reduce the impact of parameter uncertainty by minimizing the corresponding expected utility loss, $L(\{x_i\}, \{\hat{x}_i^{4F}(\beta)\})$, with respect to the trading rate β . Overall, the aim of shrinking the trading rate is to reduce the risk of taking extreme positions that may result in extreme negative outcomes.

The following proposition formulates an optimization problem whose maximizer gives the optimal shrunk trading rate of the 4-fund portfolio. Notice that we

can apply the same proposition to the 3-fund shrinkage portfolio in equation (12) simply by considering $\varsigma_2 = 0$ and $\varsigma_1 = \eta$.

Proposition 3. For the 4-fund shrinkage portfolio in equation (13), the optimal trading rate β that minimizes the expected utility loss $L(\{x_i\}, \{\hat{x}_i^{4F}(\beta)\})$ can be obtained by solving the following optimization problem:

$$(17) \quad \max_{\beta} \overbrace{V_1(x_{-1} - x^C)' \mu}^{\text{Excess return}} - \frac{1}{2} \overbrace{\left(E \left[(\hat{x}^C)' \Sigma \hat{x}^C \right] V_2 + x'_{-1} \Sigma x_{-1} V_3 + 2x'_{-1} \Sigma x^C V_4 \right)}^{\text{Variability + Trading costs}},$$

where x_{-1} is the investor's initial position, $\hat{x}^C = \varsigma_1 \hat{x}^M + \varsigma_2 \hat{x}^{\text{Min}}$ is the optimal portfolio combination between the static mean-variance portfolio and the minimum-variance portfolio,

$$(18) \quad E \left[(\hat{x}^C)' \Sigma \hat{x}^C \right] = (c/\gamma^2) \left(\varsigma_1^2 (\mu' \Sigma^{-1} \mu + (N/T)) + \varsigma_2^2 \iota' \Sigma^{-1} \iota \right) + (c/\gamma^2) (2\varsigma_1 \varsigma_2 \mu' \Sigma^{-1} \iota),$$

and elements V_i for $i = 1, \dots, 4$ are

$$(19) \quad V_1 = \frac{(1-\rho)(1-\beta)}{1-(1-\rho)(1-\beta)},$$

$$(20) \quad V_2 = \gamma \left(\frac{(1-\rho)(1-\beta)^2}{1-(1-\rho)(1-\beta)^2} - 2 \frac{(1-\rho)(1-\beta)}{1-(1-\rho)(1-\beta)} \right) + \tilde{\lambda} \frac{(1-\rho)\beta^2}{1-(1-\rho)(1-\beta)^2},$$

$$(21) \quad V_3 = \gamma \frac{(1-\rho)(1-\beta)^2}{1-(1-\rho)(1-\beta)^2} + \tilde{\lambda} \frac{(1-\rho)\beta^2}{1-(1-\rho)(1-\beta)^2},$$

$$(22) \quad V_4 = \gamma \left(\frac{(1-\rho)(1-\beta)}{1-(1-\rho)(1-\beta)} - \frac{(1-\rho)(1-\beta)^2}{1-(1-\rho)(1-\beta)^2} \right) - \tilde{\lambda} \frac{(1-\rho)\beta^2}{1-(1-\rho)(1-\beta)^2}.$$

Proposition 3 gives an optimization problem whose solution defines the investor's optimal trading rate β . As we see from expression (17), the objective is to maximize the trade-off between the expected excess return of the investor's initial portfolio with the optimal portfolio combination \hat{x}^C , and the expected portfolio variability and trading costs of the 4-fund portfolio.

To gauge the benefits from optimizing the trading rate, we compare the relative losses for the multiperiod 4-fund portfolio optimizing the trading rate as in

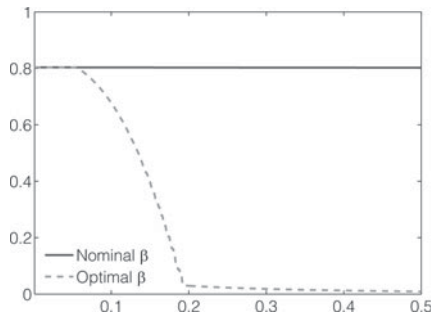
expression (17), and the multiperiod 4-fund portfolio with the nominal trading rate.

Figure 2 depicts the relative loss of the base-case investor described in Section V, assuming he or she uses a fixed estimation window with $T = 500$ observations, and population parameters μ and Σ are equal to the sample moments of the empirical data set of commodity futures described in Section V. The investor's initial portfolio is assumed to be $x_{-1} = d \times x^M$, where d is the value represented in the horizontal axis of the figure. We observe 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, we find that the relative loss of the 4-fund shrinkage portfolio can be reduced by *more than 15%* by shrinking the trading rate for the case in which the starting portfolio is $x_{-1} = 0.1 \times x^M$, and when $x_{-1} \simeq 0.5 \times x^M$, one can reduce the relative loss to almost 0. In summary, shrinking the nominal trading rate may result in a considerable reduction of the investor's expected loss, especially in those situations where the investor's initial portfolio is close to the static mean-variance portfolio.

FIGURE 2

Nominal versus Optimal 4-Fund Shrinkage Portfolios: Comparison of Relative Losses

Figure 2 depicts the investor's relative loss of our 4-fund strategy for different values of the investor's initial portfolio x_{-1} . In particular, this plot 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 $T = 500$. We define μ and Σ with the sample moments of the empirical data set of commodity futures. The investor's initial portfolio is defined as $x_{-1} = d \times x^M$, where d is the value represented on the horizontal axis. The vertical axis provides the investor's relative loss.



IV. Expanding Estimation Windows

The analysis in Sections II and III relies on the assumption that the investor uses a fixed window. We now relax this assumption by considering the case in which the investor uses expanding windows to estimate his or her portfolio; that is, the investor uses all available data at each point in time for estimation purposes. We find that for this case the expected multiperiod utility loss can no longer be separated as the product of the single-period utility loss and a multiperiod factor. Nevertheless, we conjecture an approximation to the multiperiod utility loss that allows us to estimate the optimal shrinkage intensities for the case with expanding windows. To conserve space, in the remainder of this section we briefly sketch the main steps of our analysis.

For the expanding window case, the estimated multiperiod portfolio can be written as

$$(23) \quad \hat{x}_i = (1 - \beta)^{i+1} \hat{x}_{i-1} + \beta \hat{x}_i^M,$$

where $\hat{x}_i^M = (1/\gamma) \hat{\Sigma}_i^{-1} \hat{\mu}_i$ is the estimated Markowitz portfolio at time i , and $\hat{\mu}_i$ and $\hat{\Sigma}_i$ are the estimators of the mean and covariance matrix obtained from the sample that includes all available observations up to the investment decision time i ; that is,

$$(24) \quad \begin{aligned} \hat{\mu}_i &= \frac{1}{T+i} \sum_{t=1}^{T+i} r_t, \quad \text{and} \\ \hat{\Sigma}_i &= \frac{1}{T+i-N-2} \sum_{t=1}^{T+i} (r_t - \hat{\mu}_i)^2, \end{aligned}$$

where T is the initial estimation window and i is the new available observations up to the i th investment decision stage.⁸ Note that the sample Markowitz portfolio changes over time, and thus it is not possible to use the single-period utility loss as a common factor of the investor's expected loss so that this is equal to the product between the single-period loss and a multiperiod term.

Characterizing the Multiperiod Utility Loss

To characterize the investor's expected loss with $N \geq 2$ assets, and thus to be able to optimize the shrinkage intensity for the multiperiod 3- and 4-fund portfolios, one needs to characterize expectations of the form:

$$(25) \quad B_{h,h+j} = E \left(\hat{\mu}_h' \hat{\Sigma}_h^{-1} \Sigma \hat{\Sigma}_{h+j}^{-1} \hat{\mu}_{h+j} \right),$$

where h and j are nonnegative integers. This expression is proportional to the expected out-of-sample portfolio covariance between the estimated Markowitz portfolio at time h and the estimated Markowitz portfolio at time $h+j$. This expression arises from the expected portfolio variance or trading costs of the estimated optimal trading strategy. The difficulty here is that we pre- and post-multiply the true covariance matrix with two inverse-Wishart matrices with different degrees of freedom. To the best of our knowledge, this specific problem has not been dealt with previously, and there are not available formulas to characterize this expectation.

To address this difficulty, we conjecture an approximation to the expectation in equation (25). We have tested the accuracy of our conjecture via simulations, and we find that the approximation error of our conjecture is less than 0.1%.

⁸We also considered the case of changing shrinkage intensities for an investor who uses expanding windows. We recalculate the optimal shrinkage intensities with equations (14)–(16) every time the investor has new available observations. However, this technique reduces the shrinkage intensity toward the target portfolio as time goes by, and it does not provide better out-of-sample Sharpe ratios. We do not report these results to preserve space.

Conjecture 1. Provided that $T + h > N + 4$, the expectation $B_{h,h+j}$ can be approximated as

$$(26) \quad B_{h,h+j} \approx (1 - \pi_{h,h+j})\theta + \pi_{h,h+j}B_{h,h},$$

where $\pi_{h,h+j} = (T + h - N) / (T + h + j - N)$, $\theta = \mu' \Sigma^{-1} \mu$, and the expectation $B_{h,h} = E[\hat{\mu}_h' \hat{\Sigma}_h^{-1} \Sigma \hat{\Sigma}_h^{-1} \hat{\mu}_h]$ is as given by Kan and Zhou ((2007), Sec. 2); that is, $B_{h,h} = c_h \times (\theta + N/(T + h))$, where $c_h = [(T + h - 2)(T + h - N - 2)] / [(T + h - N - 1)(T + h - N - 4)]$.

We now give some motivation for the approximation in Conjecture 1. First, we establish the bounds of equation (25) when asset returns are IID normal:

$$(27) \quad \theta \leq B_{h,h+j} < B_{h,h},$$

where $B_{h,h}$ is proportional to the expected out-of-sample portfolio variance of an estimated mean-variance portfolio with information up to time $T + h$, and θ is proportional to the true portfolio variance of the Markowitz portfolio. The true portfolio variance of the Markowitz portfolio is known to be lower than the expected out-of-sample variance of an estimated portfolio (see Kan and Zhou (2007)), and hence $\theta < B_{h,h}$.

On the other hand, we conjecture that the expected out-of-sample portfolio covariance between the estimated mean-variance portfolio constructed with information up to time $T + h$, and the estimated mean-variance portfolio constructed with information up to time $T + h + j$, is lower than the expected out-of-sample portfolio variance of an estimated mean-variance portfolio with information up to time $T + h$ (i.e., $B_{h,h+j} < B_{h,h}$). This assumption establishes the upper bound for $B_{h,h+j}$. In particular, for the specific case of $j \rightarrow \infty$, $\hat{\Sigma}_{h+j}^{-1} \hat{\mu}_{h+j} \rightarrow \Sigma^{-1} \mu$, and in turn $B_{h,h+j} = \theta < B_{h,h}$ (see Schottle and Werner (2006)).⁹

To establish the lower bound for $B_{h,h+j}$, we know that the true Markowitz portfolio provides the lowest portfolio variance, and in turn $\theta < B_{h,h+j}$. This provides the lower bound for $B_{h,h+j}$.

Therefore, it is natural that the expectation in equation (25) is between θ and $B_{h,h}$. In particular, we establish that $B_{h,h+j}$ can be characterized as a convex combination of the bounds in inequality (27), defined by parameter $\pi_{h,h+j} = (T + h - N) / (T + h + j - N)$. Notice that as j grows, the expectation approximates more toward θ because there is more information to estimate $\Sigma^{-1} \mu$. In addition, we account for the number of assets by subtracting N both in the numerator and the denominator, which acts as a smoothing term.

Given Conjecture 1, it is straightforward to characterize the expected multi-period utility loss and the optimal shrinkage intensities following the procedure used for the fixed window case described in Section III. We do not report the details to conserve space.

⁹The multivariate normal assumption implies that $\hat{\Sigma}_h$ and $\hat{\mu}_h$ are independent. As a result, when $j \rightarrow \infty$, then $B_{h,h+j} = \theta$ because $E[\hat{\Sigma}_h^{-1} \hat{\mu}_h] = \Sigma^{-1} \mu$ due to the independence and unbiasedness of $\hat{\Sigma}_h^{-1}$ and $\hat{\mu}_h$.

V. 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 or parameter uncertainty, or both. We run the comparison on simulated data sets that satisfy the assumptions of our analysis, as well as on empirical data sets. We consider both rolling as well as expanding estimation windows. Finally, we check the robustness of our results to the value of the transaction cost and absolute risk-aversion parameters, and to the estimation window length.

A. Base-Case Investor and Data Sets

We consider a base-case investor with an absolute risk-aversion parameter of $\gamma = 10^{-8}$, which corresponds to a relative risk aversion of 1 for a manager who has \$100 million to trade. GP consider an investor with a lower risk-aversion parameter, but because our investor suffers from parameter uncertainty, it is reasonable to consider a higher risk-aversion parameter. Our base-case investor has a discount factor of $\rho = 1 - \exp(-0.1/260)$, which corresponds to an annual discount of 10%. We consider an investor who is subject to quadratic transaction costs with a transaction cost parameter of $\lambda = 3 \times 10^{-7}$, as in GP. Finally, our base-case investor constructs his or her optimal trading strategy with $T = 500$ observations.¹⁰

We consider two simulated data sets with $N=25$ and 50 risky assets (N25 and N50, respectively). The advantage of using simulated data sets 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 1, 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 $[0.1, 0.5]$.¹¹ Without loss of generality, we set the return of the risk-free asset equal to 0. 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 \$364 million.¹²

We consider simulated data sets to test the out-of-sample performance of our proposed multiperiod shrinkage portfolios when the random walk assumption holds. However, it is also interesting to investigate the out-of-sample performance of our proposed strategies with empirical data sets for which this assumption may

¹⁰To estimate the shrinkage intensities, we need to estimate the population moments. To mitigate the impact of parameter uncertainty on these parameters, we use the shrinkage vector of means proposed in DeMiguel, Martín-Utrera, and Nogales (2013), and the shrinkage covariance matrix of Ledoit and Wolf (2004). Moreover, we compute the shrinkage intensities only once a month to reduce the computation time, which is important particularly for the computationally intensive expanding window approach.

¹¹For 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 prove that the investor's expected loss is proportional to $\theta = \mu' \Sigma^{-1} \mu$.

¹²To compute the trading volume of a set of assets worth \$1, we use the rule from Engle et al. (2012), who show that trading 1.59% of the daily volume implies a price change of 0.1%. Hence, for the simulated data sets we can calculate the trading volume as $1.59\% \times \text{Trading volume} \times 3 \times 10^{-7} \times 0.3/260 \times 0.5 = 0.1\%$.

fail. In particular, we consider four empirical data sets. First, we use an empirical data set formed with commodity futures (Com), similar to that used by GP (see Table 1). We collect data from those commodity futures with 3-month maturity. Some descriptive statistics and the contract multiplier for each commodity are provided in Table 1. Second, we use two equity portfolio data sets downloaded from Kenneth French's Web site:¹³ the 48 Industry portfolios (48IndP) and the Fama–French 100 portfolios (100FF) formed on size and book-to-market ratio. Finally, we also consider an individual stock data set constructed with 100 stocks (SP100) randomly selected at the beginning of each calendar year from those in the Standard & Poor's (S&P) 500 index. For the equity data sets, we download total return data and construct price change data by assuming all starting prices are equal to 1, and computing price changes from the total return data.

TABLE 1
Commodity Futures

Table 1 provides some descriptive statistics of the data from the commodity futures, as well as the contract multipliers.

| Commodity | Average Price | Volatility Price Changes | Contract Multiplier (units per contract) |
|-------------|---------------|--------------------------|--|
| 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 | 1,000 |
| RBOB crude | 88,503.62 | 2,780.74 | 42,000 |
| Natural gas | 63,553.35 | 3,439.78 | 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 |

We use daily data from July 7, 2004, until Sept. 19, 2012.¹⁴ Like GP, we focus on daily data because it is more appropriate for the investment framework we consider. To see this, note that quadratic transaction costs are typically used to model market impact costs, which occur when an investor executes a large trade that distorts market prices. Typically, investors split large trades into several smaller orders to reduce the negative impact of these price distortions (see Bertsimas and Lo (1998), Almgren and Chriss (2001)). The market impact cost literature usually focuses on high-frequency trades because low-frequency trades (low urgency to execute the portfolio) results in smaller market impact (see Almgren (2010), Engle et al. (2012)). Consequently, it seems appropriate to focus on daily data. Nevertheless, in results not reported to conserve space, we have evaluated the performance of the different portfolio policies on weekly and monthly data, and we find that our results are generally robust to the use of lower-frequency data.

¹³ See <http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/>

¹⁴ For the SP100 data set we consider daily data from July 7, 2004, until Dec. 31, 2011. We thank Grigory Vilkov for providing these data.

B. Portfolio Policies

We consider eight portfolio policies. We first consider three buy-and-hold portfolios based on single-period policies that ignore transaction costs. The first is the *sample Markowitz portfolio*, which is the portfolio of an investor who ignores transaction costs and estimation error. The second is the *single-period 2-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 toward the risk-free asset. Specifically, this portfolio can be written as

$$(28) \quad x^{S2F} = \eta \hat{x}^M,$$

where, as Kan and Zhou (2007) show, the optimal single-period shrinkage intensity η is as given by Proposition 2. The third portfolio is the *single-period 3-fund shrinkage portfolio* of an investor who ignores transaction costs but takes into account estimation error by shrinking the Markowitz portfolio toward the minimum-variance portfolio. Specifically, this portfolio can be written as

$$(29) \quad x^{S3F} = \varsigma_1 \hat{x}^M + \varsigma_2 \hat{x}^{\text{Min}},$$

where the optimal single-period combination parameters are given in Proposition 2.

We then consider five multiperiod portfolios that take transaction costs into account. The first portfolio is the *sample multiperiod portfolio* policy of an investor who takes into account transaction costs but ignores estimation error, which is given in equation (2). The second portfolio is the *multiperiod 3-fund shrinkage portfolio* of an investor who shrinks the Markowitz portfolio toward the risk-free asset, as given by Proposition 2. The third portfolio is the *multiperiod 4-fund shrinkage portfolio* of an investor who combines the Markowitz portfolio with the minimum-variance portfolio, as given by Proposition 2. The fourth portfolio is the *multiperiod 4-fund portfolio with shrunk trading rate*, which is a modified version of the multiperiod 4-fund shrinkage portfolio, where in addition the investor shrinks the trading rate by solving the optimization problem given by Proposition 3. Finally, the fifth multiperiod portfolio that we consider is the 4-fund portfolio constructed under the *expanding window approach* of Section IV.

C. Evaluation Methodology

We evaluate the out-of-sample portfolio gains for each strategy using an approach similar to DeMiguel, Garlappi, and Uppal (2009). We estimate the first seven portfolios using a rolling estimation window, in which at each point in time we use the T most recent available observations. For this rolling window approach, the length of the estimation window is constant, and hence we use the methodology introduced in Section III to construct the multiperiod shrinkage portfolios. The last portfolio is the 4-fund shrinkage portfolio estimated using an expanding window, in which at each point in time we use all observations available from $t = 1$, and we compute the shrinkage intensity using the methodology proposed in Section IV.

To account for transaction costs in the empirical analysis, we define portfolio gains net of trading costs as

$$(30) \quad R_{i+1}^k = (\hat{x}_i^k)' r_{i+1} - \tilde{\lambda} \Delta (\hat{x}_i^k)' \Sigma \Delta \hat{x}_i^k,$$

where \hat{x}_i^k denotes the estimated portfolio k at period i , r_i is the vector of price changes at the i th out-of-sample period, and Σ is the covariance matrix of asset prices.¹⁵ 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

$$(31) \quad \text{SR}^k = \frac{\bar{R}^k}{\sigma^k},$$

where

$$(32) \quad (\sigma^k)^2 = \frac{1}{L-1} \sum_{i=1}^{L-1} \left(R_{i+1}^k - \bar{R}^k \right)^2,$$

$$(33) \quad \bar{R}^k = \frac{1}{L} \sum_{i=1}^{L-1} R_{i+1}^k,$$

where L is the total number of out-of-sample periods.

We measure the statistical significance of the difference between the adjusted Sharpe ratios with the stationary bootstrap of Politis and Romano (1994), with $B = 1,000$ bootstrap samples and block size $b = 5$.¹⁶ 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 4-fund portfolio.

We consider the base-case investor described in Section V, but we also check the robustness of our results to the values of the transaction costs and absolute risk-aversion parameters, and the estimation window length. We report the results for two starting portfolios: the portfolio that is fully invested in the risk-free asset and the true Markowitz portfolio.¹⁷ We have tried other starting portfolios, such as the equal-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.

D. Performance with Rolling Estimation Window

Table 2 reports the out-of-sample Sharpe ratios of the eight portfolio policies we consider on the six different data sets, together with the p -value of the difference between the Sharpe ratio of every policy and that of the multiperiod 4-fund shrinkage portfolio. Panels A and B give the results for a starting portfolio

¹⁵For the simulated data, we use the population covariance matrix, whereas for the empirical data sets, we construct Σ with the sample estimate of the entire data set.

¹⁶We 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 $b = 5$.

¹⁷For the empirical data sets, we assume the true Markowitz portfolio is constructed with the entire sample.

TABLE 2
Sharpe Ratio Discounted with Transaction Costs

Table 2 reports the annualized out-of-sample Sharpe ratio for the different portfolio strategies that we consider. Sharpe ratios are discounted by quadratic transaction costs with $\lambda=3 \times 10^{-7}$. The numbers in parentheses are the corresponding p -values for the difference of each portfolio strategy with the 4-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 $\rho = 1 - \exp(-0.1/260)$. We consider two types of investors: one investor whose initial portfolio is fully invested in the risk-free asset, and another investor whose initial portfolio is the true Markowitz portfolio x^M . We estimate each portfolio strategy with $T = 500$ observations.

| | Data Sets | | | | | |
|---|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| | <i>N</i> = 25 | <i>N</i> = 50 | Com | 48IndP | 100FF | SP100 |
| <i>Panel A. Start from 0</i> | | | | | | |
| <i>Static Portfolio Policies</i> | | | | | | |
| Sample Markowitz | −0.266 (0.000) | −0.345 (0.000) | −0.459 (0.000) | −0.672 (0.000) | −1.435 (0.000) | −0.985 (0.000) |
| 2-fund shrinkage | 0.080 (0.000) | 0.134 (0.000) | −0.042 (0.020) | −0.100 (0.000) | −0.681 (0.000) | −0.471 (0.014) |
| 3-fund shrinkage | 0.715 (0.000) | 0.662 (0.000) | 0.587 (0.000) | −0.064 (0.000) | −0.494 (0.000) | −0.390 (0.000) |
| <i>Multiperiod Portfolio Policies</i> | | | | | | |
| Sample multiperiod | 0.150 (0.000) | 0.297 (0.008) | 0.056 (0.038) | 0.503 (0.538) | 0.209 (0.000) | −0.107 (0.110) |
| 3-fund shrinkage | 0.193 (0.000) | 0.319 (0.006) | 0.242 (0.130) | 0.525 (0.798) | 0.314 (0.022) | 0.367 (0.934) |
| 4-fund shrinkage | 0.766 (1.000) | 0.773 (1.000) | 0.893 (1.000) | 0.529 (1.000) | 0.355 (1.000) | 0.390 (1.000) |
| Trading rate shrinkage | 0.766 (0.366) | 0.773 (1.000) | 0.893 (0.602) | 0.529 (0.612) | 0.355 (1.000) | 0.390 (0.536) |
| 4-fund expanding | 0.948 (0.190) | 0.872 (0.532) | 0.803 (0.648) | 0.702 (0.496) | 0.360 (0.962) | 0.271 (0.642) |
| <i>Panel B. Start from x^M</i> | | | | | | |
| <i>Static Portfolio Policies</i> | | | | | | |
| Sample Markowitz | −0.266 (0.000) | −0.337 (0.000) | −0.452 (0.000) | −0.660 (0.000) | −1.465 (0.000) | −1.007 (0.000) |
| 3-fund shrinkage | 0.072 (0.000) | 0.131 (0.000) | −0.039 (0.010) | −0.076 (0.000) | −0.678 (0.000) | −0.470 (0.018) |
| 4-fund shrinkage | 0.714 (0.000) | 0.664 (0.000) | 0.615 (0.000) | −0.043 (0.000) | −0.490 (0.000) | −0.396 (0.000) |
| <i>Multiperiod Portfolio Policies</i> | | | | | | |
| Sample multiperiod | 0.153 (0.000) | 0.295 (0.006) | 0.052 (0.036) | 0.510 (0.524) | 0.210 (0.000) | −0.101 (0.108) |
| 3-fund shrinkage | 0.203 (0.000) | 0.309 (0.008) | 0.232 (0.080) | 0.532 (0.784) | 0.311 (0.028) | 0.370 (0.972) |
| 4-fund shrinkage | 0.774 (1.000) | 0.765 (1.000) | 0.886 (1.000) | 0.536 (1.000) | 0.351 (1.000) | 0.392 (1.000) |
| Trading rate shrinkage | 0.932 (0.070) | 0.897 (0.206) | 0.962 (0.216) | 0.766 (0.264) | 0.374 (0.712) | 0.547 (0.150) |
| 4-fund expanding | 0.957 (0.208) | 0.863 (0.486) | 0.797 (0.630) | 0.714 (0.476) | 0.356 (0.976) | 0.278 (0.694) |

that is fully invested in the risk-free asset and a starting portfolio equal to the true Markowitz portfolio, respectively.

In 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.

In comparing the shrinkage portfolios with the portfolios that ignore parameter uncertainty, we observe that shrinking helps for both the static and multiperiod

portfolios. Specifically, we find that the portfolios that shrink only the Markowitz portfolio (the single-period 2-fund shrinkage portfolio and the multiperiod 3-fund shrinkage portfolio) outperform the equivalent portfolios that ignore estimation error (the sample Markowitz portfolio and the sample multiperiod portfolio). Moreover, we find that shrinking the Markowitz portfolio toward the minimum-variance portfolio improves performance. In particular, we observe that the single-period 3-fund shrinkage portfolio and the multiperiod 4-fund shrinkage portfolio considerably outperform portfolios that shrink the Markowitz portfolio only toward the risk-free asset.

Moreover, our out-of-sample results confirm the insight from Section B that shrinking the trading rate may help when the starting portfolio is close to the true mean-variance portfolio. We see from Panel A of Table 2 that shrinking the trading rate (in addition to shrinking the Markowitz portfolio toward 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. This is not surprising, because when the investor's initial portfolio is the true Markowitz portfolio, it is optimal for the investor to avoid any trading and retain the current portfolio (i.e., $\beta = 0$). Nevertheless, the investor ignores whether his or her starting portfolio *is* the true Markowitz portfolio. The benefit of shrinking the trading rate using the rule given in expression (17) is that it takes into account whether, given the available information, the investor's starting portfolio is close to the true Markowitz portfolio or not.

Overall, the best portfolio policy is the multiperiod 4-fund shrinkage portfolio with the shrunken trading rate. This portfolio policy outperforms the multiperiod 4-fund shrinkage portfolio when the starting portfolio is close to the true Markowitz portfolio, and it performs similar to the multiperiod 4-fund shrinkage portfolio for other starting points. These two policies appreciably outperform all other policies, which shows the importance of taking into account both transaction costs and estimation error.

E. Performance with Expanding Estimation Window

We find that expanding windows generally help to improve performance. For the simulated data sets, we find that expanding windows help, which is not surprising because the simulated data sets satisfy the assumptions behind our analysis, and larger estimation windows provide more accurate estimators of the multiperiod strategies for IID data. Second, we find that expanding windows also help to improve the performance on the two empirical data sets containing data on equity portfolios (48IndP and 100FF). This is more impressive because the statistical properties of these data sets are likely to be time varying, and the expanding window approach is slower in capturing any time variation in the data. Nevertheless, we find that expanding windows help improve performance for these two data sets.

We find, however, that expanding windows do not help for the data sets containing data on commodity futures and individual stock returns (Com and SP100). We believe the reason for this may be that these data sets contain data

on individual assets (commodity futures or stocks), and thus these data sets may be more sensitive to the presence of any structural changes or time variation in the data. In turn, structural changes or time variation in the data are likely to impact the performance of the expanding window approach more negatively, because expanding windows are slower in capturing a shift in the data regime.

In summary, we believe the expanding window approach is likely to help for data sets containing equity portfolio data or relatively stationary data, and the rolling window approach is likely to perform better with data on individual assets or data that are time varying or affected by structural changes.

F. Robustness Checks

We check the robustness of our results to the transaction cost parameter λ and the estimation window length T . Note that changing the transaction cost parameter λ has the same impact on the portfolio policies as changing the absolute risk-aversion parameter γ , and thus we report the results only for the transaction cost parameter.¹⁸ We consider a high-transaction-cost scenario with $\lambda = 3 \times 10^{-6}$, which corresponds to a market with a daily trading volume of \$36.4 million, and a low-transaction-cost scenario with $\lambda = 3 \times 10^{-8}$, which corresponds to a daily trading volume of \$3,640 million for the simulated data sets (see Section V and footnote 12 to understand the relation between trading costs and trading volume). For the estimation window lengths, we consider the cases with $T = 250$ and $T = 750$.

Tables 3 and 4 report the results of the robustness checks. In general, we observe that our main insights are robust to changes in these parameters. There are substantial losses associated with ignoring both transaction costs and estimation error, and overall the best portfolio policies are the multiperiod 4-fund portfolio and the multiperiod 4-fund portfolio that in addition shrinks the trading rate.

We also observe that the static portfolio policies are very sensitive to the level of transaction costs, and their performance is particularly poor for the case with high trading costs ($\lambda = 3 \times 10^{-6}$). This is because static investors ignore transaction costs, and thus they are more vulnerable to the costs, of trading. On the other hand, multiperiod trading strategies take into account transaction costs and provide more stable portfolios with higher Sharpe ratios.

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.

¹⁸In particular, the multiperiod portfolios that are optimal for investors with $\gamma = 10^{-8}$ facing transaction costs of $\lambda = 3 \times 10^{-6}$ and $\lambda = 3 \times 10^{-8}$ are equal to those for investors with absolute risk-aversion parameters of $\gamma = 10^{-9}$ and $\gamma = 10^{-7}$, facing transaction costs of $\lambda = 3 \times 10^{-7}$, respectively. Explicitly, this means that multiplying γ by z has the same impact on the trading rate β as multiplying λ by $1/z$. Therefore, the impact of an increment/reduction on the transaction cost parameter λ is equivalent to that of a reduction/increment in γ .

TABLE 3
Sharpe Ratio: Some Robustness Checks for Different λ

Table 3 reports the annualized out-of-sample Sharpe 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 $\rho = 1 - \exp(-0.1/260)$. The investor faces quadratic transaction costs with $\lambda = 3 \times 10^{-6}$ and $\lambda = 3 \times 10^{-8}$. The numbers in parentheses are the corresponding p -values for the difference of each portfolio strategy with the 4-fund portfolio that combines the static mean-variance portfolio with the minimum-variance portfolio. We estimate each portfolio strategy with $T = 500$ observations.

| | Data Sets | | | | | |
|---|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| | <i>N</i> = 25 | <i>N</i> = 50 | Com | 48IndP | 100FF | SP100 |
| <i>Panel A. $\lambda = 3 \times 10^{-6}$</i> | | | | | | |
| <i>Static Portfolio Policies</i> | | | | | | |
| Sample Markowitz | −3.623 (0.000) | −4.020 (0.000) | −2.636 (0.000) | −2.532 (0.000) | −2.239 (0.000) | −1.002 (0.000) |
| 2-fund shrinkage | −1.099 (0.000) | −1.256 (0.000) | −1.700 (0.000) | −2.444 (0.000) | −2.012 (0.000) | −0.989 (0.000) |
| 3-fund shrinkage | 0.038 (0.000) | −0.193 (0.000) | −0.818 (0.000) | −2.382 (0.000) | −1.940 (0.000) | −1.146 (0.000) |
| <i>Multiperiod Portfolio Policies</i> | | | | | | |
| Sample multiperiod | 0.086 (0.000) | 0.219 (0.000) | −0.067 (0.014) | 0.338 (0.000) | −0.336 (0.000) | −0.552 (0.000) |
| 3-fund shrinkage | 0.187 (0.000) | 0.312 (0.010) | 0.108 (0.050) | 0.505 (0.100) | 0.020 (0.000) | 0.367 (0.764) |
| 4-fund shrinkage | 0.748 (1.000) | 0.771 (1.000) | 0.801 (1.000) | 0.528 (1.000) | 0.087 (1.000) | 0.447 (1.000) |
| Trading rate shrinkage | 0.907 (0.076) | 0.912 (0.194) | 0.917 (0.134) | 0.739 (0.274) | 0.136 (0.418) | 0.558 (0.298) |
| 4-fund expanding | 0.952 (0.182) | 0.858 (0.518) | 0.753 (0.778) | 0.696 (0.408) | 0.109 (0.990) | 0.281 (0.530) |
| <i>Panel B. $\lambda = 3 \times 10^{-8}$</i> | | | | | | |
| <i>Static Portfolio Policies</i> | | | | | | |
| Sample Markowitz | 0.141 (0.000) | 0.248 (0.008) | −0.044 (0.016) | 0.381 (0.132) | 0.089 (0.000) | −0.502 (0.016) |
| 2-fund shrinkage | 0.199 (0.004) | 0.290 (0.000) | 0.175 (0.048) | 0.396 (0.036) | 0.204 (0.000) | 0.029 (0.466) |
| 3-fund shrinkage | 0.781 (0.334) | 0.753 (0.144) | 0.841 (0.032) | 0.387 (0.014) | 0.257 (0.000) | 0.090 (0.000) |
| <i>Multiperiod Portfolio Policies</i> | | | | | | |
| Sample multiperiod | 0.179 (0.000) | 0.306 (0.010) | 0.034 (0.026) | 0.512 (0.378) | 0.341 (0.178) | −0.088 (0.304) |
| 3-fund shrinkage | 0.215 (0.000) | 0.307 (0.006) | 0.221 (0.066) | 0.476 (0.672) | 0.353 (0.056) | 0.264 (0.982) |
| 4-fund shrinkage | 0.788 (1.000) | 0.765 (1.000) | 0.884 (1.000) | 0.468 (1.000) | 0.389 (1.000) | 0.269 (1.000) |
| Trading rate shrinkage | 0.951 (0.064) | 0.874 (0.314) | 0.931 (0.368) | 0.747 (0.156) | 0.406 (0.838) | 0.496 (0.064) |
| 4-fund expanding | 0.961 (0.234) | 0.865 (0.522) | 0.796 (0.600) | 0.693 (0.374) | 0.431 (0.852) | 0.214 (0.812) |

VI. Concluding Remarks

We study the impact of parameter uncertainty in multiperiod portfolio selection with transaction costs. We provide a closed-form expression for the utility loss associated with using the plug-in approach to construct multiperiod portfolios, and we show that the investor’s expected loss decreases with trading costs, the investor’s impatience factor, and the investor’s risk-aversion parameter. In addition, we propose two complementary shrinkage approaches to mitigate the impact of parameter uncertainty: The first shrinkage approach combines the

TABLE 4
Sharpe Ratio: Some Robustness Checks for Different T

Table 4 reports the annualized out-of-sample Sharpe 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 $\rho = 1 - \exp(-0.1/260)$. The investor faces quadratic transaction costs with $\lambda = 3 \times 10^{-7}$. The numbers in parentheses are the corresponding p -values for the difference of each portfolio strategy with the 4-fund portfolio that combines the static mean-variance portfolio with the minimum-variance portfolio.

| | Data Sets | | | | | |
|---------------------------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| | $N = 25$ | $N = 50$ | Com | 48IndP | 100FF | SP100 |
| <i>Panel A. $T = 250$</i> | | | | | | |
| <i>Static Portfolio Policies</i> | | | | | | |
| Sample Markowitz | -1.126 (0.000) | -1.458 (0.000) | -0.766 (0.000) | -1.543 (0.000) | -4.815 (0.000) | -1.456 (0.000) |
| 2-fund shrinkage | -0.127 (0.000) | -0.044 (0.000) | -0.138 (0.006) | -0.989 (0.000) | -3.318 (0.000) | -1.225 (0.000) |
| 3-fund shrinkage | 0.459 (0.000) | 0.535 (0.000) | 0.046 (0.000) | -0.883 (0.000) | -2.448 (0.000) | -0.882 (0.000) |
| <i>Multiperiod Trading Strategies</i> | | | | | | |
| Sample multiperiod | 0.055 (0.000) | 0.218 (0.000) | 0.635 (0.762) | 0.315 (0.000) | -0.790 (0.000) | -0.835 (0.000) |
| 3-fund shrinkage | 0.159 (0.000) | 0.334 (0.006) | 0.490 (0.488) | 0.545 (0.558) | -0.008 (0.000) | 0.052 (0.672) |
| 4-fund shrinkage | 0.640 (1.000) | 0.746 (1.000) | 0.710 (1.000) | 0.563 (1.000) | 0.164 (1.000) | 0.162 (1.000) |
| Trading rate shrinkage | 0.765 (0.076) | 0.935 (0.094) | 0.688 (0.278) | 0.567 (0.952) | 0.168 (0.042) | 0.241 (0.560) |
| 4-fund expanding | 0.938 (0.102) | 0.929 (0.322) | 0.829 (0.562) | 0.392 (0.560) | 0.304 (0.702) | 0.366 (0.538) |
| <i>Panel B. $T = 750$</i> | | | | | | |
| <i>Static Portfolio Policies</i> | | | | | | |
| Sample Markowitz | -0.023 (0.000) | 0.207 (0.000) | -0.156 (0.156) | -0.098 (0.000) | -0.570 (0.000) | -0.998 (0.000) |
| 2-fund shrinkage | 0.080 (0.000) | 0.440 (0.006) | -0.186 (0.068) | 0.247 (0.000) | -0.071 (0.000) | -0.526 (0.012) |
| 3-fund shrinkage | 0.709 (0.004) | 0.844 (0.000) | 0.307 (0.044) | 0.290 (0.000) | -0.011 (0.000) | -0.316 (0.000) |
| <i>Multiperiod Trading Strategies</i> | | | | | | |
| Sample multiperiod | 0.216 (0.000) | 0.545 (0.038) | 0.049 (0.324) | 0.608 (0.964) | 0.398 (0.118) | -0.145 (0.198) |
| 3-fund shrinkage | 0.167 (0.000) | 0.568 (0.048) | -0.120 (0.142) | 0.592 (0.262) | 0.456 (0.796) | 0.140 (0.628) |
| 4-fund shrinkage | 0.755 (1.000) | 0.917 (1.000) | 0.468 (1.000) | 0.612 (1.000) | 0.462 (1.000) | 0.301 (1.000) |
| Trading rate shrinkage | 0.829 (0.478) | 0.987 (0.586) | 0.566 (0.036) | 0.938 (0.320) | 0.524 (0.620) | 0.454 (0.196) |
| 4-fund expanding | 0.924 (0.170) | 0.851 (0.664) | 0.685 (0.272) | 0.506 (0.572) | 0.387 (0.798) | 0.169 (0.558) |

Markowitz portfolio with a target portfolio that is less sensitive to estimation error, and the second shrinkage approach reduces the investor trading rate in order to avoid the risk of taking extreme positions that may result in extreme negative outcomes. Finally, we show analytically and empirically the superior performance of those portfolios that account for the impact of both parameter uncertainty and transaction costs.

Appendix. Proofs

Proof of Proposition 1. We first write the investor's expected loss as the difference between the utility of an investor who knows the population vector of means μ and the covariance

matrix Σ , and the expected utility of an investor who replaces μ and Σ with their sample estimates:

$$(A-1) \quad L(\{x_i\}, \{\hat{x}_i\}) = \sum_{i=0}^{\infty} (1-\rho)^{i+1} \left\{ x_i' \mu - \frac{\gamma}{2} x_i' \Sigma x_i - \frac{\tilde{\lambda}}{2} \Delta x_i' \Sigma \Delta x_i \right. \\ \left. - E \left[\hat{x}_i' \mu - \frac{\gamma}{2} \hat{x}_i' \Sigma \hat{x}_i - \frac{\tilde{\lambda}}{2} \Delta \hat{x}_i' \Sigma \Delta \hat{x}_i \right] \right\}.$$

Because the estimated optimal trading strategy is unbiased, one can write the investor's expected loss as the sum of the following terms:

$$(A-2) \quad L(\{x_i\}, \{\hat{x}_i\}) = \underbrace{\frac{\gamma}{2} \sum_{i=0}^{\infty} (1-\rho)^{i+1} E [\hat{x}_i' \Sigma \hat{x}_i - x_i' \Sigma x_i]}_{\text{Multiperiod mean-variance loss}} \\ + \underbrace{\frac{\tilde{\lambda}}{2} \sum_{i=0}^{\infty} (1-\rho)^{i+1} E [\Delta \hat{x}_i' \Sigma \Delta \hat{x}_i - \Delta x_i' \Sigma \Delta x_i]}_{\text{Transaction cost loss}}.$$

From equation (A-2), we can observe that the multiperiod expected loss is the sum of a multiperiod mean-variance term and a transaction cost term. Now, we plug the estimated investor's optimal strategy into equation (A-2) in order to obtain a simplified expression of the investor's expected loss. First, notice that those elements defined as linear functions of the sample Markowitz portfolio disappear due to the unbiasedness of the estimator. Second, we use the following expressions of the estimated multiperiod portfolio:

$$(A-3) \quad \begin{aligned} \hat{x}_i &= (1-\beta)^{i+1} x_{-1} + \beta \xi_i \hat{x}^M \quad \text{and} \\ \Delta \hat{x}_i &= \phi_i x_{-1} + \beta(1-\beta)^i \hat{x}^M, \end{aligned}$$

where $\xi_i = \sum_{j=0}^i (1-\beta)^j$, $\phi_i = ((1-\beta)^{i+1} - (1-\beta)^i)$, and x_{-1} is the investor's initial portfolio. We estimate the investor's multiperiod strategy with a fixed window T . Then, after some straightforward manipulations, we obtain that the investor's expected loss is the product between the expected loss of a static investor and another term that accounts for the multiperiod effects of the mean-variance utility and the transaction costs:

$$(A-4) \quad L(\{x_i\}, \{\hat{x}_i\}) = \frac{1}{2\gamma} \left(E [\hat{\mu}' \hat{\Sigma}^{-1} \Sigma \hat{\Sigma}^{-1} \hat{\mu}] - \theta \right) \\ \times \sum_{i=0}^{\infty} (1-\rho)^{i+1} [AV_i + AC_i],$$

where $\theta = \mu' \Sigma^{-1} \mu$. The multiperiod term is the sum of two terms. The first term, $AV_i = \beta^2 \xi_i^2$ captures the utility losses from the accumulated portfolio variability. In particular, β^2 can be understood as the proportion of the expected variability of a static mean-variance investor that affects a multiperiod investor at every stage i , and ξ_i^2 determines the impact of the discounted expected variability of all trades up to stage i . The second term $AC_i = (\tilde{\lambda}/\gamma) \beta^2 (1-\beta)^{2i}$ captures the losses from the accumulated trading costs. In particular, $\beta^2 (1-\beta)^{2i}$ is the impact on transaction costs of trading toward the static mean-variance portfolio, at stage i . We observe that as i becomes larger, the impact on transaction costs is lower because the optimal multiperiod portfolio is closer to the static mean-variance portfolio, and as a result the optimal multiperiod strategy requires less trading. Then, we can substitute $(1/2\gamma)(E[\hat{\mu}' \hat{\Sigma}^{-1} \Sigma \hat{\Sigma}^{-1} \hat{\mu}] - \theta)$ with $L_1(x^M, \hat{x}^M)$, and using the fact that

ξ_i can be written as $\sum_{j=0}^i (1-\beta)^j = (1 - (1-\beta)^{i+1})/\beta$ and the properties of infinite geometric series, we can express the accumulated variability and transaction costs as

$$(A-5) \quad f_{mv} = \beta^2 \sum_{i=0}^{\infty} (1-\rho)^{i+1} \xi_i^2$$

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

$$(A-6) \quad f_{tc} = \sum_{i=0}^{\infty} (1-\rho)^{i+1} \beta^2 \frac{\tilde{\lambda}}{\gamma} (1-\beta)^{2i} = \frac{\lambda}{\gamma} \frac{\beta^2}{1-(1-\rho)(1-\beta)^2}.$$

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

$$(A-7) \quad L(\{x_i\}, \{\hat{x}_i\}) = L_1(x^M, \hat{x}^M) \times [f_{mv} + f_{tc}].$$

Proof of Proposition 2. We now prove that the optimal combination parameters of multiperiod portfolios coincide with the optimal combination parameters in the static framework. For expository reasons, we focus on the proof of the multiperiod 4-fund portfolio, but it is straightforward to do the analysis for the multiperiod 3-fund portfolio. First, let us define the investor's initial portfolio as x_{-1} . Then, we can write the investor's 4-fund portfolio \hat{x}_i^{4F} and the difference $\Delta \hat{x}_i^{4F}$ as

$$(A-8) \quad \hat{x}_i^{4F} = (1-\beta)^{i+1} x_{-1} + \beta \xi_i \hat{x}^C,$$

$$(A-9) \quad \Delta \hat{x}_i = \phi_i x_{-1} + \beta(1-\beta)^i \hat{x}^C,$$

where $\hat{x}^C = (\varsigma_1 \hat{x}^M + \varsigma_2 \hat{x}^{\text{Min}})$, $\xi_i = \sum_{j=0}^i (1-\beta)^j$ and $\phi = ((1-\beta)^{i+1} - (1-\beta)^i)$. Plugging equations (A-8) and (A-9) into the investor's expected utility, we obtain:

$$(A-10) \quad E \left[\sum_{i=0}^{\infty} (1-\rho)^{i+1} \left\{ (1-\beta)^{i+1} x'_{-1} \mu + \beta \xi_i (\hat{x}^C)' \mu \right. \right. \\ \left. \left. - \frac{\gamma}{2} \left((1-\beta)^{2i+2} x'_{-1} \Sigma x_{-1} + \beta^2 \xi_i^2 (\hat{x}^C)' \Sigma \hat{x}^C + 2\beta(1-\beta)^{i+1} \xi_i x'_{-1} \Sigma \hat{x}^C \right) \right. \right. \\ \left. \left. - \frac{\tilde{\lambda}}{2} \left(\phi_i^2 x'_{-1} \Sigma x_{-1} + \beta^2 (1-\beta)^{2i} (\hat{x}^C)' \Sigma \hat{x}^C + 2\phi_i \beta (1-\beta)^i x'_{-1} \Sigma \hat{x}^C \right) \right\} \right].$$

Now, using the properties of geometric series as in Proposition 1, we obtain that the investor's expected utility is defined as follows:

$$(A-11) \quad \tau_1 (x_{-1} - x^C)' \mu + \tau_5 x^C' \mu \\ - \frac{\gamma}{2} \left\{ \tau_2 x'_{-1} \Sigma x_{-1} + (\tau_5 + \tau_2 - 2\tau_1) E((\hat{x}^C)' \Sigma \hat{x}^C) + 2(\tau_1 - \tau_2) x_{-1}' \Sigma x^C \right\} \\ - \frac{\tilde{\lambda}}{2} \left\{ \beta^2 \tau_3 x'_{-1} \Sigma x_{-1} + \beta^2 \tau_3 E((\hat{x}^C)' \Sigma \hat{x}^C) + 2\beta(\tau_4 - \tau_3) x'_{-1} \Sigma x^C \right\},$$

where

$$\tau_1 = \frac{(1-\rho)(1-\beta)}{1-(1-\rho)(1-\beta)^2}, \quad \tau_2 = \frac{(1-\rho)(1-\beta)^2}{1-(1-\rho)(1-\beta)^2}, \\ \tau_3 = \frac{(1-\rho)}{1-(1-\rho)(1-\beta)^2}, \quad \tau_4 = \frac{(1-\rho)(1-\beta)}{1-(1-\rho)(1-\beta)^2}, \quad \text{and} \\ \tau_5 = \frac{1-\rho}{\rho}.$$

We can obtain the optimal value of ς_1 by developing the first-order conditions that maximize the investor's expected utility with respect to ς_1 . This is

$$(A-12) \quad \varsigma_1 = \frac{(x^M)' \mu}{\gamma E[(\hat{x}^M)' \Sigma \hat{x}^M]} \frac{W_1}{W_2} - \frac{x'_{-1} \Sigma x^M}{\gamma E[(\hat{x}^M)' \Sigma \hat{x}^M]} \frac{W_3}{W_2} - \varsigma_2 \frac{E[(\hat{x}^M)' \Sigma \hat{x}^{\text{Min}}]}{E[(\hat{x}^M)' \Sigma \hat{x}^M]},$$

where $W_1 = \tau_5 - \tau_1$, $W_2 = (\tau_5 + \tau_2 - 2\tau_1) + (\tilde{\lambda}/\gamma)\beta^2\tau_3$, and $W_3 = \gamma(\tau_1 - \tau_2) + \tilde{\lambda}\beta(\tau_4 - \tau_3)$. We numerically verify that $W_1/W_2 = 1$ and $W_3 = 0$, so that the optimal parameter ς_1 takes the following expression:

$$(A-13) \quad \varsigma_1 = \frac{(x^M)' \mu}{\gamma E[(\hat{x}^M)' \Sigma \hat{x}^M]} - \varsigma_2 \frac{E[(\hat{x}^M)' \Sigma \hat{x}^{\text{Min}}]}{E[(\hat{x}^M)' \Sigma \hat{x}^M]}.$$

Similarly, we obtain the optimal value of ς_2 by developing the first-order conditions that maximize the investor's expected utility. Thus, the optimal ς_2 takes the following form:

$$(A-14) \quad \varsigma_2 = \frac{(x^{\text{Min}})' \mu}{\gamma E[(\hat{x}^{\text{Min}})' \Sigma \hat{x}^{\text{Min}}]} - \varsigma_1 \frac{E[(\hat{x}^M)' \Sigma \hat{x}^{\text{Min}}]}{E[(\hat{x}^{\text{Min}})' \Sigma \hat{x}^{\text{Min}}]}.$$

Therefore, one can solve the system given by equations (A-13) and (A-14) 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))

$$(A-15) \quad \varsigma_1 = c^{-1} \frac{\Psi^2}{\Psi^2 + \frac{N}{T}},$$

$$(A-16) \quad \varsigma_2 = c^{-1} \frac{\frac{N}{T}}{\Psi^2 + \frac{N}{T}} \times \frac{\mu' \Sigma^{-1} \iota}{\iota' \Sigma^{-1} \iota},$$

where $c = [(T-2)(T-N-2)]/[(T-N-1)(T-N-4)]$, and $\Psi^2 = \mu' \Sigma^{-1} \mu - (\mu' \Sigma^{-1} \iota)^2 / (\iota' \Sigma^{-1} \iota) > 0$. Therefore, one can obtain the optimal value of η by setting $\varsigma_2 = 0$ in equation (A-13). As a result, the optimal value of η is (see Kan and Zhou (2007))

$$(A-17) \quad \eta = \frac{(x^M)' \mu}{\gamma E[(\hat{x}^M)' \Sigma \hat{x}^M]} = c^{-1} \frac{\mu' \Sigma^{-1} \mu}{\mu' \Sigma^{-1} \mu + \frac{N}{T}}.$$

Our results of the optimal combination parameters are slightly different with respect to those of Kan and Zhou (2007) because we consider the sample covariance matrix that yields an unbiased estimator of the inverse covariance matrix.

Proof of Corollary 1. We know from Proposition 2 that the optimal combination parameters coincide with the optimal combination parameters of the static 3-fund portfolio of Kan and Zhou (2007). Then, we can show that it is optimal to shrink the static mean-variance portfolio if the derivative of the (single-period) investor's 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 that it is optimal to have $\eta < 1$ when

$$(A-18) \quad (x^M)' \mu < \gamma E[(\hat{x}^M)' \Sigma \hat{x}^M] \Rightarrow$$

$$(A-19) \quad \mu' \Sigma^{-1} \mu < E[\hat{\mu}' \hat{\Sigma}^{-1} \Sigma \hat{\Sigma}^{-1} \hat{\mu}].$$

Because we assume that price changes are multivariate normal, $\hat{\mu}$ and $\hat{\Sigma}$ are independent.¹⁹ Then, to characterize the expectations in inequality (A-19), we use the property for the expected value of quadratic forms of a random vector x with mean μ and covariance matrix Σ . This property says that $E[x'Ax] = \mu'A\mu + \text{trace}(A\Sigma)$, with A being a definite positive matrix. Moreover, we can characterize the expectation of $\hat{\Sigma}^{-1}\Sigma\hat{\Sigma}^{-1}$ using the identities of inverse-Wishart matrices described in Haff (1979). In particular, we have that $E[\hat{\Sigma}^{-1}\Sigma\hat{\Sigma}^{-1}] = c \times \Sigma^{-1}$, where $c = [(T - N - 2)(T - 2)] / [(T - N - 1)(T - N - 4)]$. Then, we obtain that $\eta < 1$ if $\theta < c(\theta + N/T)$. Because $c > 1$, we observe that it is always optimal to shrink the static mean-variance portfolio.

Now, we show that it is optimal to shrink the sample Markowitz portfolio toward the sample minimum-variance portfolio. If we take derivatives of the (single-period) investor's expected utility with respect to parameter ς_2 , and then set $\varsigma_2 = 0$, this derivative is positive (and in turn it is optimal to have $\varsigma_2 > 0$) if

$$(A-20) \quad (x^{\text{Min}})' \mu > \gamma \varsigma_1 E[(\hat{x}^{\text{M}})' \Sigma \hat{x}^{\text{Min}}] \Rightarrow$$

$$(A-21) \quad \iota' \Sigma^{-1} \mu > \varsigma_1 E[\hat{\mu}' \hat{\Sigma}^{-1} \Sigma \hat{\Sigma}^{-1} \iota].$$

Now, we can again characterize these expectations by using the identities of inverse-Wishart matrices derived in Haff (1979). Thus, 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 $\Psi^2 = (\mu - \mu_g \iota)' \Sigma^{-1} (\mu - \mu_g \iota)$, where $\mu_g = (\iota' \Sigma^{-1} \mu) / (\iota' \Sigma^{-1} \iota)$. As a result, Ψ^2 is nonnegative, and the relation $1 > \varsigma_1 c$ always holds. Notice that ς_2 cannot be negative because then 1 has to be lower than $\varsigma_1 c$, which is not possible because $\Psi^2 > 0$.

Proof of Proposition 3. We can prove Proposition 3 by writing the investor's expected utility as in the proof of Proposition 2, which is the expected utility for an investor using the multiperiod 4-fund portfolio. Moreover, we remove from the investor's expected utility all the elements τ_5 because the trading rate does not appear in that term, and in turn it would not affect the optimization problem. Then, we make some straightforward arrangements to obtain the following expression:

$$(A-22) \quad V_1(x_{-1} - x^C)' \mu - \frac{1}{2} \left(E[(\hat{x}^C)' \Sigma \hat{x}^C] V_2 + x'_{-1} \Sigma x_{-1} V_3 + x'_{-1} \Sigma x^C V_4 \right),$$

where V_i is defined as in equations (19)–(22) in the paper.

We see that equation (A-22) is formed first by the expected excess return of the investor's initial portfolio with the optimal portfolio combination \hat{x}^C , and second by a term that represents the expected variability and trading costs of the 4-fund portfolio. Now, we can characterize $E[(\hat{x}^C)' \Sigma \hat{x}^C]$ using the property for the expected value of a quadratic form of a random vector, as defined in the proof of Corollary 1. We also use the identities for inverse-Wishart matrices to characterize $E[\hat{\Sigma}^{-1} \Sigma \hat{\Sigma}^{-1}]$ as in the proof of Corollary 1. Then, we obtain

$$\begin{aligned} (A-23) \quad E[(\hat{x}^C)' \Sigma \hat{x}^C] &= \frac{1}{\gamma^2} E \left[\varsigma_1^2 \hat{\mu}' \hat{\Sigma}^{-1} \Sigma \hat{\Sigma}^{-1} \hat{\mu} + \varsigma_2^2 \iota' \hat{\Sigma}^{-1} \Sigma \hat{\Sigma}^{-1} \iota \right. \\ &\quad \left. + 2\varsigma_1 \varsigma_2 \hat{\mu}' \hat{\Sigma}^{-1} \Sigma \hat{\Sigma}^{-1} \iota \right] \\ &= \frac{c}{\gamma^2} \left(\varsigma_1^2 \left(\mu' \Sigma^{-1} \mu + \frac{N}{T} \right) + \varsigma_2^2 \iota' \Sigma^{-1} \iota + 2\varsigma_1 \varsigma_2 \mu' \Sigma^{-1} \iota \right), \end{aligned}$$

¹⁹Due to the multivariate normality assumption, $\hat{\mu}$ is normally distributed as $N(\mu, \Sigma/T)$, and $\hat{\Sigma}$ follows a Wishart distribution of the form $\mathcal{W}(\Sigma/(T - N - 2), T - 1)$.

where $c = [(T - N - 2)(T - 2)] / [(T - N - 1)(T - N - 4)]$. Notice that the analysis can be easily extended for the case of the sample multiperiod portfolio or the multiperiod 3-fund shrinkage portfolio.

References

- Almgren, R. "Execution Costs." In *Encyclopedia of Quantitative Finance*. Chichester, England: John Wiley & Sons (2010), 612–616.
- Almgren, R., and N. Chriss. "Optimal Execution of Portfolio Transactions." *Journal of Risk*, 3 (2001), 5–39.
- Bertsimas, D., and A. W. Lo. "Optimal Control of Execution Costs." *Journal of Financial Markets*, 1 (1998), 1–50.
- Best, M. J., and R. R. Grauer. "Positively Weighted Minimum-Variance Portfolios and the Structure of Asset Expected Returns." *Journal of Financial and Quantitative Analysis*, 27 (1992), 513–537.
- Brown, S. J. "The Portfolio Choice Problem: Comparison of Certainty Equivalence and Optimal Bayes Portfolios." *Communications in Statistics—Simulation and Computation*, 7 (1978), 321–334.
- Constantinides, G. M. "Multiperiod Consumption and Investment Behavior with Convex Transactions Costs." *Management Science*, 25 (1979), 1127–1137.
- Cornuejols, G., and R. Tutuncu. *Optimization Methods in Finance*. Cambridge, MA: Cambridge University Press (2007).
- Davis, M. H. A., and A. R. Norman. "Portfolio Selection with Transaction Costs." *Mathematics of Operations Research*, 15 (1990), 676–713.
- DeMiguel, V.; L. Garlappi; F. J. Nogales; and R. Uppal. "A Generalized Approach to Portfolio Optimization: Improving Performance by Constraining Portfolio Norms." *Management Science*, 55 (2009), 798–812.
- DeMiguel, V.; L. Garlappi; and R. Uppal. "Optimal versus Naïve Diversification: How Inefficient Is the $1/N$ Portfolio Strategy?" *Review of Financial Studies*, 22 (2009), 1915–1953.
- DeMiguel, V.; A. Martín-Utrera; and F. J. Nogales. "Size Matters: Optimal Calibration of Shrinkage Estimators for Portfolio Selection." *Journal of Banking and Finance*, 37 (2013), 3018–3034.
- Engle, R., and R. Ferstenberg. "Execution Risk." *Journal of Portfolio Management*, 33 (2007), 34–44.
- Engle, R.; R. Ferstenberg; and J. Russell. "Measuring and Modeling Execution Cost and Risk." *Journal of Portfolio Management*, 38 (2012), 14–28.
- Garlappi, L.; R. Uppal; and T. Wang. "Portfolio Selection with Parameter and Model Uncertainty: A Multi-Prior Approach." *Review of Financial Studies*, 20 (2007), 41–81.
- Garleanu, N., and L. H. Pedersen. "Dynamic Trading with Predictable Returns and Transaction Costs." *Journal of Finance*, 68 (2013), 2309–2340.
- Goldfarb, D., and G. Iyengar. "Robust Portfolio Selection Problems." *Mathematics of Operations Research*, 28 (2003), 1–38.
- Greenwood, R. "Short- and Long-Term Demand Curves for Stocks: Theory and Evidence on the Dynamics of Arbitrage." *Journal of Financial Economics*, 75 (2005), 607–649.
- Haff, L. R. "An Identity for the Wishart Distribution with Applications." *Journal of Multivariate Analysis*, 9 (1979), 531–544.
- Jagannathan, R., and T. Ma. "Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps." *Journal of Finance*, 58 (2003), 1651–1684.
- Kan, R., and G. Zhou. "Optimal Portfolio Choice with Parameter Uncertainty." *Journal of Financial and Quantitative Analysis*, 42 (2007), 621–656.
- Kirby, C., and B. Ostdiek. "It's All in the Timing: Simple Active Portfolio Strategies That Outperform Naïve Diversification." *Journal of Financial and Quantitative Analysis*, 47 (2012), 437–467.
- Klein, R. W., and V. S. Bawa. "The Effect of Estimation Risk on Optimal Portfolio Choice." *Journal of Financial Economics*, 3 (1976), 215–231.
- Kyle, A. S. "Continuous Auctions and Insider Trading." *Econometrica*, 53 (1985), 1315–1335.
- Ledoit, O., and M. Wolf. "A Well-Conditioned Estimator for Large-Dimensional Covariance Matrices." *Journal of Multivariate Analysis*, 88 (2004), 365–411.
- Ledoit, O., and M. Wolf. "Robust Performance Hypothesis Testing with the Sharpe Ratio." *Journal of Empirical Finance*, 15 (2008), 850–859.
- Lillo, F. J.; J. D. Farmer; and R. N. Mantegna. "Master Curve for Price-Impact Function." *Nature*, 421 (2003), 129–130.
- Liu, H. "Optimal Consumption and Investment with Transaction Costs and Multiple Risky Assets." *Journal of Finance*, 59 (2004), 289–338.

- MacKinlay, A. C., and L. Pastor. "Asset Pricing Models: Implications for Expected Returns and Portfolio Selection." *Review of Financial Studies*, 13 (2000), 883–916.
- Markowitz, H. "Portfolio Selection." *Journal of Finance*, 7 (1952), 77–91.
- Pástor, L. "Portfolio Selection and Asset Pricing Models." *Journal of Finance*, 55 (2000), 179–223.
- Pástor, L., and R. F. Stambaugh. "Comparing Asset Pricing Models: An Investment Perspective." *Journal of Financial Economics*, 56 (2000), 335–381.
- Politis, D., and J. Romano. "The Stationary Bootstrap." *Journal of the American Statistical Association*, 89 (1994), 1303–1313.
- Rustem, B.; R. G. Becker; and W. Marty. "Robust Min-Max Portfolio Strategies for Rival Forecast and Risk Scenarios." *Journal of Economic Dynamics and Control*, 24 (2000), 1591–1621.
- Schottle, K., and R. Werner. "Towards Reliable Efficient Frontiers." *Journal of Asset Management*, 7 (2006), 128–141.
- Tu, J., and G. Zhou. "Incorporating Economic Objectives into Bayesian Priors: Portfolio Choice under Parameter Uncertainty." *Journal of Financial and Quantitative Analysis*, 45 (2010), 959–986.
- Tutuncu, R. H., and M. Koeing. "Robust Asset Allocation." *Annals of Operations Research*, 132 (2004), 157–187.
- Wang, Z. "A Shrinkage Approach to Model Uncertainty and Asset Allocation." *Review of Financial Studies*, 18 (2005), 673–705.