Quarterly Review of Economics and Finance 100 (2025) 101962

Contents lists available at ScienceDirect

Quarterly Review of Economics and Finance

journal homepage: www.elsevier.com/locate/qref

Markowitz portfolios under transaction costs

Olivier Ledoit, Michael Wolf *

Department of Economics, University of Zurich, CH-8032 Zurich, Switzerland

ARTICLE INFO

JEL classification:
C13
G11

Keywords:
Covariance matrix estimation
Mean-variance efficiency
Multivariate GARCH
Portfolio selection
Transaction costs

ABSTRACT

Markowitz portfolio selection is a cornerstone in finance, in academia as well as in the industry. Most academic studies either ignore transaction costs or account for them in a way that is both unrealistic and suboptimal by (i) assuming transaction costs to be constant across stocks and (ii) ignoring them at the portfolio-selection state and simply paying them after the fact. Our paper proposes a method to fix both shortcomings. As we show, if transaction costs are accounted for (properly) at the portfolio-selection stage, net performance in terms of the Sharpe ratio often increases, in particular for high-turnover strategies.

1. Introduction

Markowitz (1952) portfolios are a cornerstone of finance, both in the academic literature and in the asset-management world. Since the two main input parameters, the vector of expected returns and the covariance matrix of the returns in a universe of assets, are unknown in practice, the true (or ideal) portfolios are infeasible and have to be estimated instead. To this end one most commonly applies the plug-in method: use the estimated vector of expected returns and the estimated covariance matrix in place of the true quantities.

When several methods to estimate a given portfolio under consideration are available, it is generally of interest to determine which method is best. In particular when researchers come up with a new method, they generally want to demonstrate that their method is better than existing methods. One way to answer this question is via Monte Carlo studies. Unlike in real life, in a Monte Carlo study the true portfolio is known and so one can determine how close (on average) a given method gets to the truth. But any Monte Carlo study is based on a data generating process (DGP) chosen by the researchers, which leaves room to tweak this process in their favor. Therefore, the more common way to answer the question is via backtest exercises: One applies the various methods to a set of real data in a realistic manner (meaning that at any given point only data prior to that point are used to estimate the portfolio) over a sufficiently long period, which results in a series of (pseudo) out-of-sample (oos) returns for each method. Then the various oos return series are used for performance evaluation.

Which criterion to use in the evaluation depends on the portfolio under consideration. In the case of the Global Minimum Variance (GMV) portfolio, arguably, the most relevant criterion is risk, measured by the sample variance (or, equivalently, the sample standard deviation) of the oos return series. The respective sample numbers then provide a ranking of the various methods (with smaller numbers resulting in a higher ranking). In addition, one can use hypothesis testing to examine whether one method significantly outperforms another one; for example, see Ledoit and Wolf (2011). On the other hand, a general Markowitz portfolio not only seeks a small risk but also a high reward. Whereas the former is quantified by the standard deviation of the portfolio return, the latter is quantified by its expectation. In such cases, the preferred criterion is the reward-to-risk ratio, that is, the expected return divided by the standard deviation. When the returns are taken in excess of the risk-free rate, this ratio is known as the Sharpe ratio; otherwise, the ratio is often called the information ratio. In a backtest exercise one then uses the sample average of the oos return series divided by the sample standard deviation. Again, one can use hypothesis tests to study whether one method significantly outperforms another one; for example, see Ledoit and Wolf (2008).

When the Sharpe ratio is used as the performance criterion, the question becomes whether to account for transaction costs or not. In our view, the answer is "it depends". Although this distinction is rarely made in the literature, there can be two goals in performance evaluation.

The first goal, which one could term accuracy, is to evaluate the quality of estimated inputs in plug-in methods, namely of the estimated vector of expected returns and of the estimated covariance matrix; the idea being that the higher the quality of the estimated inputs, the higher will be the oos Sharpe ratio. Therefore, in this context, a backtest serves as an indirect way of evaluating the quality of estimated

* Correspondence to: Zürichbergstrasse 14, CH-8032 Zurich, Switzerland.
E-mail addresses: olivier.ledoit@econ.uzh.ch (O. Ledoit), michael.wolf@econ.uzh.ch (M. Wolf).

https://doi.org/10.1016/j.qref.2025.101962
Received 17 January 2024; Received in revised form 18 September 2024; Accepted 3 November 2024
Available online 12 January 2025

inputs. (The direct way, by comparing estimated inputs to true inputs, is not feasible, since the true inputs are not observable.) For this goal, transaction costs should *not* be taken into account, since doing so would muddy the waters. For example, consider the case where one wants to compare two estimators of the covariance matrix (holding the estimator of the vector of expected returns fixed). It could well be that one of the estimators, although less precise, leads to lower turnover and thus to a higher Sharpe ratio *after* transaction costs.

The second goal, which one could term efficiency (shortcut for the mean–variance efficiency of a dynamically-rebalanced portfolio), is to evaluate actual investment strategies. Of course, for this goal, transaction costs should be taken into account. Arguably, an asset manager does not care about the quality of estimated inputs *per se* as long as they result in superior performance, which for him/her always means after transaction costs. Many academic studies take transaction costs into account, be it as the leading objective or as a robustness check. The problem is that the majority of such studies do so in a non-realistic and suboptimal way: (i) they assume transaction costs to be constant across stocks[1] and (ii) they do not take transaction costs into account at the portfolio-selection stage; for example, see DeMiguel et al. (2009), Neely et al. (2014), Ao et al. (2019), Kan et al. (2022), and Chen et al. (2024).

Concerning (i), assuming that stocks have the same transaction costs makes life easier but contradicts reality. Therefore, we incorporate this feature, using a realistic off-the-shelf transaction-cost model from the literature. Concerning (ii), few portfolio managers in the real world ignore transaction costs at the portfolio-selection stage and simply pay them after the fact. Doing so can be expected to result in worse net performance (that is, performance after transaction costs), at least for high-turnover strategies. Therefore, we propose a way to account for transaction costs at the portfolio-selection stage.

The remainder of the paper is organized as follows. Section 2 specifies the portfolio formulations of interest and how we account for transaction costs at the portfolio-selection stage. Section 3 details our backtesting technology. Section 4 presents the results based on real stock-return data. Section 5 concludes. An appendix contains some additional details and tables.

2. Portfolios

2.1. The starting point

Our starting point is the standard Markowitz portfolio-selection problem in its textbook version. There are N assets in the universe with corresponding return vector $r := (r_1, \dots, r_N)'$. Denote the expected value and the covariance matrix of r by

$$\mu := \mathbb{E}(r) \quad \text{and} \quad \Sigma := \mathsf{Cov}(r) .$$

The portfolio-selection problem in its textbook version is then formulated as

$$\min_w w'\Sigma w \tag{2.1}$$

$$\text{subject to} \quad w'\mu \ge b \text{ , and} \tag{2.2}$$

$$w'\mathbb{1} = 1 \text{ ,} \tag{2.3}$$

where b is a selected target expected return and $\mathbb{1}$ denotes a conformable vector of ones.[2] The problem has the following analytical solution:

$$w = c_1 \Sigma^{-1}\mathbb{1} + c_2\Sigma^{-1}\mu \text{ ,} \tag{2.4}$$

where $c_1 := \dfrac{C - bB}{AC - B^2}$ and $c_2 := \dfrac{bA - B}{AC - B^2}$, (2.5)

$$\text{with } A := \mathbb{1}'\Sigma^{-1}\mathbb{1} \text{ ,} \quad B := \mathbb{1}'\Sigma^{-1}\mu \text{ ,} \quad \text{and} \quad C := \mu'\Sigma^{-1}\mu \text{ .} \tag{2.6}$$

Since μ and Σ are unknown in practice, this solution is not feasible. But a feasible solution can be obtained by replacing μ with an estimator $\hat{\mu}$ and Σ with an estimator $\hat{\Sigma}$ in the analytical formulas: this is the plug-in method.

Although of academic interest, the plug-in version of the solution (2.4)–(2.6) is rarely used in real life. On the one hand, portfolio managers generally are interested in further constraints in addition to (2.2) and (2.3), such as gross-exposure, maximum(-absolute)-position, industry-exposure, and factor-exposure constraints; of these, as a leading case, we will consider a gross-exposure constraint in our analysis below. On the other hand, the formulation (2.1)–(2.2) completely ignores transaction costs. In real life, the portfolio manager already holds a current portfolio and the question is how to best update (or turn over) the portfolio. But updating (or turning over) a portfolio involves transaction costs, which are stock-specific. Of course, one can simply ignore transaction costs at the portfolio-selection stage, find the best portfolio given one's list of constraints, and then pay the resulting transaction costs after the fact. As can be expected theoretically, and as we will confirm empirically below, doing so is often inferior to taking transaction costs into account at the portfolio-selection stage already.

2.2. Accounting for transaction costs

There is no unique way to take transaction costs into account. One way would be to include an upper bound on such costs as an additional constraint; the question then becomes what that amount should be. Instead, we opt for adding a penalty term, defined as a multiple of the transaction cost, to the objective function (2.1). The question then becomes what the multiplier should be. We feel that it is more natural to address the latter question than the former.

To start out, the more general problem formulation needed now must address the fact that one is moving form a holding portfolio to a new portfolio, so the concept of time needs to be introduced. In our empirical analysis below, we will use daily stock data, which is the norm in literature; but, as also is the norm, we do not update the portfolio on a daily basis but only every 21 trading days, which corresponds to our notion of a "month".[3] As a result, the notation must distinguish between (a) return time and (b) portfolio-selection (or rebalancing) time. Return days are indexed by $t \in \{1, \dots, T\}$, where T denotes the sample size; portfolio-selection dates are indexed by $h \in \{1, \dots, H\}$, where H denotes the number of months. The connection is that t_h denotes the (return) day corresponding to portfolio-selection date h; in particular, $t_{h+1} = t_h + 21$.

Consider a portfolio-selection date $h > 1$. The current vector of holding-portfolio weights is given by

$$w^*_{h-1} := (w^*_{h-1,1}, \dots, w^*_{h-1,N_h})' .$$

This vector specifies the portfolio at the end of the holding period $h-1$ and differs from w_{h-1}, which specifies the portfolio at the beginning of the holding period. The two vectors differ because of the different price evolutions of the various stocks during the holding period; see (3.4) for a description how to obtain w^*_h for a general holding period h. Finally, N_h denotes the size of the combined investment universe at portfolio-selection date h. This universe is given as the union of all stocks in the current holding portfolio and all stocks to be included in the new (or upcoming) portfolio: At any given date $h > 1$, one has to allow for some stocks leaving the portfolio and others entering the portfolio. So even if the number of stocks in the holding portfolio remains constant (denoted by N), as will be the case for us, $N_h \ge N$ generally varies

[1] To be safe, this means transaction costs are assumed to be constant across stocks at any given point in time, but might be allowed to change, and in particular decrease, over time.

[2] One can also formulate the problem in its dual form: maximizing the expected return subject to an upper bound on the variance of the portfolio.

[3] So a 'month' generally does not correspond to a calendar month.

over time. Of course, in case $N_h > N$, there will be (at least) $N_h - N$ positions in w_{h-1}, w^*_{h-1}, and w_h that are equal to zero.

The next ingredient is a vector of stock-specific transaction costs[4] at portfolio-selection date h:

$$c_h := (c_{h,1}, \dots, c_{h,N_h})' \ .$$

The unit of the transaction costs must match the unit of the stock returns. Say a given transaction cost is 10 basis points (bps). If stock returns are raw returns, then this number should be expressed as 0.001; on the other hand, if the stock returns are in percent, then this number should be expressed as 0.1 instead.[5] When updating (or turning over) the portfolio from w^*_{h-1} to w_h the total transaction cost "per dollar" is then given by

$$\tau_h := c_h'|w_h - w^*_{h-1}| := \sum_{i=1}^{N_h} c_{h,i}|w_{h,i} - w^*_{h-1,i}| \ . \tag{2.7}$$

Therefore, τ_h is a relative rather than an absolute cost. The absolute cost is obtained by multiplying the relative cost with the current net-asset value (NAV) of the portfolio. As an example, take the case where τ_h equals 5 bps. If the NAV equals one hundred million dollars, then the cost of turning over the portfolio is fifty thousand dollars; if the NAV equals one billion dollars, the cost is half a million dollars.

The advantage of expressing the transaction cost per dollar instead of in absolute terms is that the penalty parameter on the transaction cost, which is about to be introduced, can also be chosen per dollar and thus does not have to be adapted to the current value of the portfolio, which changes over time.

To facilitate the upcoming methodology it will be convenient to partition the combined investment universe at portfolio-selection date h into two sets: the one containing the stocks in the new portfolio (of size N) and the one containing the stocks leaving the holding portfolio (of size M_h) with $M_h \geq 0$ and $N_h = N + M_h$. For further convenience, the stocks will be (re-)ordered such that the first N will effectively constitute the new portfolio and the last M_h will leave the holding portfolio. In this way, the total transaction cost (per dollar) can be expressed as the sum of (i) a variable cost (corresponding to the first N stocks) and (ii) a fixed cost (corresponding to the last M_h stocks):

$$\tau_h = \tau_h^{\text{var}} + \tau_h^{\text{fix}} := \sum_{i=1}^{N} c_{h,i}|w_{h,i} - w^*_{h-1,i}| + \sum_{i=N+1}^{N_h} c_{h,i}|w^*_{h-1,i}| \ . \tag{2.8}$$

The latter cost (which is zero in case $N_h = N$ and positive otherwise) is obviously fixed, since it does not depend on how the weights for the stocks in the new portfolio are selected. On the other hand, this selection affects the former cost, which is therefore variable.

The two final ingredients are $\hat{\mu}_h$ and $\hat{\Sigma}_h$, the estimators of μ_h and Σ_h which, in slight abuse of notation, are defined as

$$\mu_h := \mathbb{E}(r^{\text{new}}_{t_h}) \quad \text{and} \quad \Sigma_h := \text{Cov}(r^{\text{new}}_{t_h}) \quad \text{with} \quad r^{\text{new}}_{t_h} := (r^{\text{new}}_{t_h,1}, \dots, r^{\text{new}}_{t_h,N})' \ .$$

Clearly only the returns $r^{\text{new}}_{t_h}$ of the stocks in the new investment universe are relevant, in the sense that we need estimators of the expectation and the covariance matrix of the vector of returns; the stocks leaving the holding portfolio are irrelevant. At this point, we assume nothing about the nature of $\hat{\mu}_h$ and $\hat{\Sigma}_h$; specific choices will be discussed in the empirical analysis below.

With the definition of τ_h^{fix} as in (2.8), our formulation of the portfolio-selection problem, both accounting for transaction costs and incorporating a gross-exposure constraint, is then given by:

$$\min_{w \in \mathbb{R}^N} w'\hat{\Sigma}_h w + \lambda \cdot \sum_{i=1}^{N} c_{h,i}|w_i - w^*_{h-1,i}| \tag{2.9}$$

$$\text{subject to} \quad w'\hat{\mu}_h \geq b_h \ , \tag{2.10}$$

$$w'\mathbb{1} = 1 \ , \text{ and} \tag{2.11}$$

$$\|w\|_1 \leq \kappa \ . \tag{2.12}$$

The constant λ in (2.9) is the penalty parameter on the transaction cost per dollar. Alternatively, and equivalently, one could formulate the objective function (2.9) as

$$\min_{w \in \mathbb{R}^N} w'\hat{\Sigma}_h w + \lambda \cdot \left(\sum_{i=1}^{N} c_{h,i}|w_i - w^*_{h-1,i}| + \tau_h^{\text{fix}} \right) ,$$

noting that τ_h^{fix} does not depend on w. Furthermore, our formulation allows for the bound b_h on the expected return in (2.10) to vary with time.[6] Finally, $\|\cdot\|_1$ in (2.12) denotes the L_1 norm of a vector, that is,

$$\|w\|_1 := \sum_{i=1}^{N} |w_i| \ .$$

Therefore, (2.12) represents a gross-exposure constraint where $\kappa \geq 1$ expresses the upper bound.[7] In the extreme case $\kappa = 1$, the portfolio is long only, meaning that $w_i \geq 0$ for all i. The choice $\kappa = 1.6$ corresponds to a 130-30 portfolio whereas the choice $\kappa = 2$ corresponds to a 150-50 portfolio; such portfolios are becoming ever more popular in the financial industry. The choice $\kappa = \infty$ effectively removes the constraint (2.12) from the portfolio formulation.

Denote the solution of (2.9)–(2.12) by $\hat{w}$. Then $w_h := (\hat{w}_1, \dots, \hat{w}_N, 0, \dots, 0)'$, where the last $M_h \geq 0$ positions equal zero. To be careful, when portfolio-selection date $h+1$ comes, w^*_h shall be re-ordered relative to w_h such that the last $M_{h+1} \geq 0$ positions correspond to the stocks that will leave the holding portfolio at date $h+1$.

It should be pointed out that the problem (2.9)–(2.12) cannot be solved analytically, that is, there is no formula for $\hat{w}$ (unless $\lambda = 0$ and $\kappa = \infty$). However, the problem is of convex nature and can be solved with off-the-shelf numerical-optimization software coming under the header "convex solver". Even for $N = 1,000$, solving the problem only takes a matter of seconds (in the single digits) with modern computers.

2.3. Variations on the theme

Many other portfolio formulations based on the plug-in method can be considered, where the formulations can be either "smaller" or "larger".

To go smaller, one can drop the expected-return constraint (2.10). Doing so has the advantage that one does not require an estimator of the expected return vector μ_h, which is not trivial to come by with. Note that by now, an entire academic industry, if the pun may be forgiven, is dedicated to finding such estimators, which sometimes are called "factors" or "(return-predictive) signals"; for example, see Green et al. (2013) and Harvey et al. (2016). If in addition the gross-exposure constraint (2.12) is dropped, one arrives at the global minimum variance (GMV) portfolio accounting for transaction costs; setting $\lambda = 0$ in (2.9) gives the standard 'textbook' GMV portfolio (in its feasible version based on the plug-in method). But, again, the standard version is not of much interest in real life. On the one hand, the gross exposure can be anything and vary quite a bit from one portfolio-selection date to another, which for many portfolio managers would be problematic or not even be permitted. On the other hand, transaction costs are not accounted for. As for the general Markowitz portfolio, our empirical study will compare the performance of the textbook formulation with formulations accounting for transactions costs and/or incorporating a gross-exposure constraint.

[4] Our methodology is agnostic about how stock-specific transaction costs are obtained (or modeled); we discuss a specific proposal to this end in Section 3.2.

[5] Note that 100 bps correspond to one percent.

[6] This is to allow for some variation in expected market conditions.

[7] One could, in principle, also allow κ to vary with time but doing so is uncommon.

To go 'larger', as would be typical in real life, one can add further constraints after (2.12), such as a maximum(-absolute)-position constraint (putting a common upper bound on $|w_i|$ to avoid overexposure to any given stock), industry-exposure constraints, and factor-exposure constraints. Importantly, adding any such constraints (in their common forms) preserves the convexity of the optimization problem, so that $\hat{w}$ can still be easily found with off-the-shelf numerical optimization software.

2.4. Estimation of input parameters

The feasible portfolio formulation (2.9)–(2.12) requires estimates of two input parameters, μ_h and Σ_h. The formulation is high-level in the sense that nothing is said about the nature of the estimators $\hat{\mu}_h$ and $\hat{\Sigma}_h$, which must be based on a (finite) history of past observations.

It is common knowledge which estimators *not* to use (without further modifications to the resulting solution $\hat{w}$): namely, the sample analogs, meaning the vector of sample means (for $\hat{\mu}_h$) and the sample covariance matrix (for $\hat{\Sigma}_h$). Although unbiased, and maximum likelihood under a normality assumption, in settings where the number of stocks and the number of past observations are of similar magnitude (which is the setting most portfolio managers face) the sample analogs contain too much estimation error. Consequently, portfolios $\hat{w}$ based on them perform poorly out of sample and are sometimes called "estimation error maximizers", a term coined by Michaud (1989).

Therefore, it is of interest to find improved estimators of the two input parameters in practice. Thankfully, this task can be considered as a division of labor. Some people aim to find improved estimators of the vector of expected returns; for example, see again Green et al. (2013) and Harvey et al. (2016). Other people aim to find improved estimators of the covariance matrix; for example, see Ledoit and Wolf (2022a) and the references therein.[8]

Finally, we have a methodology in place that can be safely used also when the number of assets is large; by which we mean up to a couple thousand of assets.[9] On the one hand, the methodology proposed in Sections 2.2–2.3 allows for the selection of portfolios within a second or less once input parameters have been estimated; on the other hand, improved estimators of μ and Σ result in favorable oos performance, as will be seen in our empirical study of Section 4.

2.5. Non-Markowitz paradigms

Our proposal is firmly placed within the paradigm of Markowitz portfolio selection. Arguably, this is still the most popular such paradigm in both academia and the industry. Needless to say, many alternative paradigms exist as well. They can differ from the Markowitz paradigm in any possible direction and dimension. We are not privy to what is used in the industry but even the paradigms proposed in the literature are too numerous to list them all. To point out just a few examples, see Kan and Zhou (2007), Brandt et al. (2009), Frahm and Memmel (2010), Tu and Zhou (2011), Kolm et al. (2014), López de Prado (2016, 2020), Ao et al. (2019), Dixon et al. (2020), Kan et al. (2022), and Kelly and Xiu (2023).

There are three potential problems regarding methods that belong to alternative portfolio-selection paradigms, which are routinely swept under the academic carpet but greatly matter to portfolio managers in the industry: First, can the method handle large investment universes and the case $N > T$? Second, can the method account for transaction costs at the portfolio-selection stage? Third, can the method incorporate additional constraints, such as a gross-exposure constraint, factor-exposure constraints, or a maximum(-absolute)-position constraint? Appendix A discusses two exemplary non-Markowitz paradigms in some detail.

2.6. Related literature

The literature on accounting for transaction costs at the portfolio-selection stage is surprisingly sparse. Nevertheless, there are some previous proposals which we can put in perspective.

Liu (2004) extends the concept of a "no-trade zone", which had been long established in the literature for a single risky asset, to multiple risky assets. For tractability reasons, he specifies a model where the decision whether or not to trade a given stock is independent across stocks. We are likewise sympathetic to the concept of a no-trade zone: On any given selection date, our method may refuse to trade a substantial fraction of the universe, especially for large values of the penalty parameter λ. However, the method does so in such a way that all trade/no-trade decisions are taken jointly, and every stock can influence the others, which is more realistic and expected to yield better results.

Apart from transaction costs, three things matter: (return-predictive) signals, risk, and optimization. There is a division of labor between Gârleanu and Pedersen (2013) and our own paper. Whereas these authors analyze the impact of transaction costs on the signal mix while taking the risk model and optimizer for granted, we take the signals for granted while analyzing the impact of transaction costs on the optimization process within the context of an accurate and up-to-date high-dimensional risk model. Their main insight, that short-lived signals must be downweighted relative to long-lived signals due to round-trip transaction-cost drag, is completely sound, but it has little bearing on our approach.

Brandt et al. (2009) also employ the concept of a no-trade zone, which they choose as a hyper-sphere of radius k around the current vector of holding-portfolio weights w^*_{h-1}. If the vector of the target-portfolio weights, which is obtained by selecting an "optimal" portfolio ignoring transaction costs, lies outside the no-trade zone, one trades right onto the boundary of the zone in direction of the target portfolio; this corresponds to a convex combination of the holding portfolio and the target portfolio. There are three concerns regarding this proposal: (i) it does not reflect non-constant transaction costs across stocks, that is, it might be more appropriate to use a hyper-ellipse instead of a hyper-sphere as the shape of the no-trade zone; (ii) it is not clear how to pick the size of the no-trade zone, that is, the radius k in case of a hyper-sphere; and (iii) it is not clear how to handle changing portfolio universes, since some trading needs to take place if the universe changes.

DeMiguel et al. (2020) investigate how transaction costs change the number of characteristics that are jointly significant for an investor's optimal portfolio and, hence, how they change the dimension of the cross-section of stock returns. Specifically, they show that transaction costs noticeably increase the number of significant characteristics and that using a larger number of characteristics in parametric portfolios reduces turnover and transaction costs, leading to superior out-of-sample performance net of transaction costs. This approach is fundamentally different from ours in that it accounts for transaction costs implicitly rather than explicitly and does not allow portfolio managers to dial down (or up) turnover to their liking.

[8] Improved estimation of the covariance matrix is huge field in the academic literature and this not the place to give a comprehensive overview. Our point is that improved estimators (relative to the sample analog) exist and we take the liberty of highlighting our own ones in this regard without necessarily claiming that they are necessarily better than any other such estimators.

[9] Obviously, the term "large" is up to a certain amount of interpretation and even in this decade some people in the literature seem to interpret it as "more than two".

3. Backtesting methodology

3.1. Data and universe rules

We download daily stock return data from the Center for Research in Security Prices (CRSP) starting on 01-Jan-1995 and ending on 31-Dec-2022. We restrict attention to stocks from the NYSE, AMEX, and NASDAQ exchanges. Daily risk-free rates come from the supplemental series of the CRSP US Treasury Database via the Wharton Research Data Services portal.[10]

For simplicity, we adopt the common convention that 21 consecutive trading days constitute one 'month'. The (pseudo) out-of-sample period ranges from 19-Jan-2000 through 30-Dec-2022, resulting in a total of $H = 275$ months (or $T = 5,775$ days). All portfolios are updated monthly.[11] We index the portfolio-selection dates by $h = 1, \dots, 275$.

We now detail our rules for determining the investment universe at any given portfolio-selection date h. For a stock to be eligible, it must satisfy five criteria. The first criterion is not feasible in practice but common in academic analyses (and, basically, never kicks in anyhow for large stocks): The stock must have a complete return future over the next 21 days. Second, the day when the stock started, that is, the day where we see for the first time a non-missing value for the return of the stock must be at least 1260 days in the past. Third, over the history of the past 1260 days, there must be a most 2.5% missing values; any such missing value is then replaced with the return of the S&P 500 index on that day. Fourth, there must be no pairs of stocks with respective return series over the past 1260 days that have a sample correlation exceeding 0.95. To ensure this criterion, we compute the sample correlation matrix of all the stock satisfying the first three criteria. If the largest correlation does not exceed 0.95, we keep all stocks. Otherwise, we identify the pair with the largest correlation and remove the stock with the smaller market value of the two at date h; then we repeat the exercise until there are no more correlations exceeding 0.95. There is a final, fifth, criterion, which is not based on daily returns. For reasons that will become obvious below we also need price data in open, high, low, and close (OHLC) form. For a stock to be eligible, it needs to have all four pieces of information available for at least 240 out of the past 252 days. Then, at any date h, the universe is comprised of the $N = 1,000$ largest eligible stocks in terms of their average market value over the past month (that is, over the 21 days most recent days in the past).

3.2. Modeling transaction costs

The goal of our paper is to develop a strategy that delivers desirable portfolio returns net of transaction costs. Importantly, we are interested in a strategy that will work well now, that is, from here on going forward. We are not interested in a strategy that would have worked well in the past, that is, over the last previous thirty years, say.

Consequently, we will use "historical" stock returns in combination with "current" transaction costs, the idea being that historical stock returns are (hopefully) representative of future ones and that current transactions are also representative of future ones, at least in the short- and medium-term future.

Apart from wanting to use current transaction costs, we also want to be realistic and use stock-specific transactions costs. To this end, we use the model of Briere et al. (2020), which works by modeling stock-specific bid–ask spreads and then taking transaction costs to be a constant fraction (across stocks) of these spreads. In particular, they propose to model the log bid–ask spread of stock i at day t as a linear function of past log volatility. The problem is that volatility is unobserved even in hindsight, so in an initial step a feasible "version" of volatility is taken to be the Garman and Klass (1980) intraday estimate using past OHLC price data; see Equation [11.15] of Briere et al. (2020):

$$\sigma_{t,i}^{\mathrm{GK}} := \sqrt{\frac{1}{252}\sum_{s=1}^{252}\frac{1}{2}\log\left(\frac{H_{t-s,i}}{L_{t-s},i}\right)^2 - (2\log 2 - 1)\log\left(\frac{C_{t-s,i}}{O_{t-s,i}}\right)^2} \ . \tag{3.1}$$

If the complete OHLC information is not available for all past 252 days, we base the estimate on the corresponding formula applied only to the days when the complete information is available, of which there must be at least 240 days by our criteria above.

Based on a regression analysis the authors then propose the following linear approximation for the log bid–ask spread of stock i on day t (see their Table 11.3):

$$\log bas_{t,i} \approx -4.137 + 0.777\log\sigma_{t,i}^{\mathrm{GK}} \ ,$$

which results in the approximated transaction cost of stock i on day t as follows:

$$c_{t,i} := \frac{1}{2}\exp\left(-4.137 + 0.777\log\sigma_{t,i}^{\mathrm{GK}}\right) \ . \tag{3.2}$$

Here, using a constant multiplier across stocks, we equate transaction cost to one half of the bid–ask spread, which is a common convention. Briere et al. (2020) actually propose 0.4 instead of 0.5 as the multiplier (see their Table 11.1) but we prefer to stick to the commonly used and somewhat more conservative value of 0.5.

Finally, with some abuse of notation, the approximation we use in the problem formulation (2.9)–(2.12) is given by $c_{h,i} := c_{t_h,i}$, where it should be recalled that t_h denotes the day t that corresponds to portfolio-selection date h. Although an approximation only itself, it is certainly more realistic than assuming a constant trading cost across stocks, which is still the norm in the related academic literature. In our experience, the transaction-cost model of Briere et al. (2020) is a realistic place-holder by default, but any applied researcher who already holds a strong preference for some alternative model is welcome to use it instead, as our methodology is not tied to a specific transaction-cost model. For some alternative models, see Hasbrouck (2009), Novy-Marx and Velikov (2016), Abdi and Ranaldo (2017), Ardia et al. (2024), and the references therein.

To get an idea what typical trading costs look like using approximation (3.2), we carry out the following exercise: For every portfolio-selection date $h = 1, \dots, 275$, take the vector of the $N = 1,000$ approximate trading costs corresponding to the stocks in the new portfolio. Based on this vector, compute the minimum, the first decile, the first quartile, the median, the third quartile, the ninth decile, and the maximum. Then, for each of these seven quantiles $q \in \{0, 0.1, 0.25, 0.5, 0.75, 0.9, 1\}$, compute the average over the 275 portfolio-selection dates h; note that, with some abuse of terminology, we call the minimum the 0 quantile and the maximum the 1 quantile. The results are reported in Table 3.1. As can be seen, for example, the median transaction cost is 3.6 bps on average. It can also be seen that there is considerable variation across transaction costs, as the 0.9 quantile is, in terms of the average, almost twice as large as the 0.1 quantile.

3.3. Computing returns net of transaction costs

We need to be careful about how we compute oos-ntc returns, (where the acronym oos-ntc stands for "out-of-sample and net(-of)-transaction-costs"), and we need to state our method clearly. The problem is that we have daily portfolio returns but update the portfolio

[10] The unique series identifier is 2000061, corresponding to "RISKFREE2 (MTH/DLY)". The maturity/ rebalancing label is "CRSP Risk Free – 4 week (Nominal)".

[11] Monthly updating is common practice to avoid an unreasonable amount of turnover and thus transaction costs. During a month, from one day to the next, we hold number of shares fixed rather than portfolio weights; in this way, there are no transactions during a month.

Table 3.1
Average quantiles of approximate trading costs; the unit is 1 basis point. At each portfolio-selection date $h = 1, \dots, 275$, the seven listed quantiles are computed from the vector of $N = 1{,}000$ approximate trading costs (of the stocks in the new investment universe) according to formula Eq. (3.2). Then, for each quantile, the average over the 275 portfolio-selection dates is reported.

Quantile	0.0	0.1	0.25	0.5	0.75	0.9	1.0
Average	1.7	2.7	3.1	3.6	4.3	5.2	10.0

only every month, that is, every 21 days. In terms of the average oos-ntc return it would not matter if we paid the transaction cost in full at once. But doing so would unduly affect the standard deviation of the returns (which also enters our performance criteria below), since only one out of twenty-one returns would take the full hit; so instead we spread out evenly the incurred transaction cost over all 21 days in the upcoming holding period, that is, over the stretch $\{t_h, \dots, t_h + 20\}$. The cleanest way to do this is in terms of the net-asset value (NAV) series of the portfolio.

At the first portfolio-selection date $h = 1$, we (somewhat cavalierly) ignore transaction costs, since at this date the portfolio is not updated but formed for the first time.[12] Based on an (arbitrary) starting value X_0, one invests amount X_0 according to weight vector w_1 at the beginning of day 1.

Since the value of X_0 is irrelevant for our purposes, we set it to $X_0 := 1$ for convenience. The initial stretch of the NAV series, namely $\{X_t\}_{t=1}^{21}$, is determined by letting the portfolio run for 21 days, keeping the number of shares (rather than the portfolio weights) fixed over time:

$$\forall t = 1, \dots, 21 \qquad X_t = w_1' \, \Pi_{s=1}^{t} (1 + r_s) \ ,$$

where multiplication of two vectors is understood to be element-wise, that is,

$$(a_1, \dots, a_N)' \cdot (b_1, \dots, b_N)' := (a_1 b_1, \dots, a_N b_N)' \ .$$

By the principle of recursion, to define the NAV series over the entire oos period, it now is sufficient to specify our "recipe" for a generic date $h > 1$. Denote by X_t the NAV of the portfolio at the end of day t where, of course, "net" also means "net of transaction costs". The transaction cost incurred at portfolio-selection date h is given by $\tau_h \cdot X_{t_{h-1}}$ with τ_h given by (2.7). In an initial step, we derive a no-transaction-cost asset-value series $\{\tilde{X}_{t_h}, \dots, \tilde{X}_{t_h+20}\}$ as follows: Invest amount X_{t_h} according to weight vector w_h at the beginning of day t_h and let the portfolio run for 21 days, holding the number of shares (rather then the portfolio weights) fixed:

$$\forall j = 0, \dots, 20 \qquad \tilde{X}_{t_h+j} = X_{t_h-1} \cdot w_h' \, \Pi_{s=0}^{j} (1 + r_{t_h+s}) \ .$$

The NAV series is then defined as

$$\forall j = 0, \dots, 20 \qquad X_{t_h+j} := \tilde{X}_{t_h+j} - (\tau_h \cdot X_{t_h-1}) \frac{j+1}{21} \ .$$

This convention corresponds to paying off the transaction cost incurred evenly over the 21 days in the upcoming holding period.[13]

In this way, a NAV series $\{X_t\}_{t=0}^{T}$ is obtained, with $X_0 = 1$ by our above convention. From the NAV series one then backs out the oos-ntc returns as follows:

$$\forall t = 1, \dots, T \qquad x_t := \frac{X_t}{X_{t-1}} - 1 \ . \tag{3.3}$$

The series $\{x_t\}_{t=1}^{T}$ then forms the basis for performance evaluation.

Last but not least, in this context we can also detail how w_h^* is obtained, the vector of portfolio weights at the end of holding period h:

$$w_h^* \propto w_h \cdot \Pi_{s=0}^{20} (1 + r_{t_h+s}) \ , \tag{3.4}$$

where the symbol $\propto$ stands for "proportional to"; since the weights need to some up to one, specifying the proportional weights is obviously sufficient.

4. Empirical study

4.1. Investment strategies

We consider four portfolio formulations:

- **GMV:** (2.9) and (2.11).
- **GMV-130-30**: in addition (2.12) with $\kappa = 1.6$.
- **Marko:** (2.9)–(2.11).
- **Marko-130-30:** in addition (2.12) with $\kappa = 1.6$.

To judge the benefit of accounting for transaction costs at the portfolio-selection stage, the portfolios GMV and Marko are of leading interest to us. Since portfolio managers in the industry often face additional constraints, we also include the portfolios GMV-130-30 and Marko-130-30 as a real-life robustness check.

4.2. Estimators in the horse race

We consider two estimators $\hat{\Sigma}_h$ of the covariance matrix Σ_h:

- **NL**: the static nonlinear shrinkage estimator of Ledoit and Wolf (2022b, Section 4.5), termed QIS (Quadratic-Inverse Shrinkage) estimator.
- **DCC-NL**: the dynamic multivariate GARCH estimator of Engle et al. (2019). Note that we use the QIS estimator to estimate the correlation targeting matrix C instead of the QuEST function; see their Section 3.3.

As per the suggestion of Engle et al. (2019), we use the previous 1260 days as input to the dynamic DCC-NL estimator, which roughly corresponds to using 5 years of past data. On the other, we only use 630 days as input to the static NL estimator, which roughly corresponds to using 2.5 years of past data. The reason is that that a static estimator changes (much) more slowly compared to a dynamic estimator if the same number of days is used as input. Another way to look at this issue is that a dynamic estimator down-weights the distant past relative to the recent past via its model structure whereas a static estimator gives the same weight to all days in the estimation window. Therefore, using a shorter estimation window can be seen as a crude weighting scheme pertaining to the longer window comprising 1260 days: give equal weight to all 630 days in the short window of "recent past" and zero to all 630 days in the remaining window of "distant past".

As a simple-minded benchmark for the GMV portfolios, we include the equally-weighted portfolio, called **EW** which *de facto* assumes that the covariance matrix is proportional to the identity matrix.

The portfolios Marko and Marko-130-30 need an estimator $\hat{\mu}_h$ in addition. For simplicity and reproducibility, we use the well-known momentum factor (or simply momentum for short) of Jegadeesh and Titman (1993). For a given portfolio-selection date h and a given stock, the momentum is the geometric average of the previous 252 returns on the stock but excluding the most recent 21 returns; in other words, one uses the geometric average over the previous year but excluding the previous month. Collecting the individual momentums of all the N stocks contained in the new universe then yields $\hat{\mu}_h$.

As a simple-minded benchmark for the Marko portfolios we include the equally-weighted portfolio of the top-quintile stocks (according to momentum). This portfolio is obtained by sorting the stocks, from lowest to highest, according to their momentum and then putting equal weight on all the stocks in the top 20%, that is, in the top quintile.

[12] We are not investigating how to build a portfolio once from scratch but how to update a portfolio repeatedly through time.

[13] It is tacitly assumed here that no interest will be charged by paying off the cost over the 21 upcoming days instead of paying it off in full at once on date h.

We call this portfolio **EW-TQ**. We then use the value of b_h implied for the EW-TQ portfolio as input in the constraint (2.11), that is, b_h is the average momentum of the stocks in the EW-TQ portfolio.[14]

4.3. Performance measurement

We report the following three out-of-sample performance measures for each scenario, where the acronym oos-ntc stands for "out-of-sample and net-of-transaction-costs" and the corresponding returns are defined in (3.3):

- **AV:** We compute the average of the 8568 oos-ntc returns in excess of the risk-free rate and then multiply by 252 to annualize.
- **SD:** We compute the standard deviation of the 8568 oos-ntc returns in excess of the risk-free rate and then multiply by $\sqrt{252}$ to annualize.
- **SR:** We compute the (annualized) Sharpe ratio as the ratio AV/SD.

As stated in the introduction, there are two possible goals for a backtest exercise: (i) the evaluation of the quality of estimated input parameters (like $\hat{\Sigma}_h$) and (ii) the evaluation of actual investment strategies. The goal in this paper is the second one and, therefore, for all four portfolio formulations considered, the leading performance criterion of interest is SR.[15]

In addition, we report the following two summary statistics pertaining to the weight vectors:

- **TO:** Average monthly turnover defined as $\frac{1}{274}\sum_{h=2}^{275}\|w_{h-1}^{*}-w_h\|_1$.
- **GE:** Average gross exposure defined as $\frac{1}{275}\sum_{h=1}^{275}\|w_h\|_1$.

4.4. Transaction-cost penalty

Our problem formulations account for transaction costs via the penalty parameter λ in (2.9). Unfortunately, there is no theory dictating an optimal choice of λ. Instead we try out different values in a certain grid and report the various measures for each choice. The grid considered is

$$\lambda \in \{0, 2.5, 5, 7.5, 10, 15, 20, 50\}\ ,$$

where the choice $\lambda = 0$ corresponds to the *status quo* in the academic literature: ignore transaction costs at the portfolio-selection stage and simply pay them after the fact; as we aim to show, doing so is typically sub-optimal with respect to SR.

Detailed results can be found in Tables B.1–B.4 in the appendix. In order to mitigate the effects of "cherry picking" and "Monday-morning quarterbacking", we limit ourselves to picking a common "good" value of λ across the four different portfolios considered and also the two covariance matrix estimators considered. The following general findings can be observed.

- Increasing the penalty parameter λ reduces TO. This is a strict monotonic relation which holds in every scenario, as can be anticipated from theory.
- Although the relation is not strictly monotonic, increasing λ tends to increase AV. This makes sense, since penalizing transaction costs more reduces them, which benefits the average return net of transaction costs.

Table 4.1
Annualized performance measures (in percent) for various estimators of the GMV and GMV-130-30 portfolios. AV stands for average; SD stands for standard deviation; and SR stands for Sharpe ratio. All measures are based on 5,775 daily out-of-sample returns net of transaction costs and are in excess of the risk-free rate. Furthermore, TO stands for average monthly turnover and GE stands for average gross exposure.

	EW	NL_0	$NL_{7.5}$	$DCC-NL_0$	$DCC-NL_{7.5}$
			GMV Portfolio		
AV	10.41	6.75	7.71	7.63	7.69
SD	21.64	9.89	9.86	8.14	7.20
SR	0.48	0.68	0.78	0.94	1.07
TO	0.10	1.30	0.38	2.84	1.08
GE	1.00	4.41	4.24	3.14	3.51
			GMV-130-30 Portfolio		
AV	10.41	8.84	8.61	6.82	7.06
SD	21.64	10.51	10.59	7.80	7.46
SR	0.48	0.84	0.81	0.87	0.95
TO	0.10	0.45	0.24	1.76	1.11
GE	1.00	1.60	1.60	1.60	1.60

Table 4.2
Analogous to Table 4.1 but now for the Marko and Marko-130-30 portfolios.

	$EW-TQ$	NL_0	$NL_{7.5}$	$DCC-NL_0$	$DCC-NL_{7.5}$
			Marko Portfolio		
AV	9.51	6.88	7.82	6.75	7.16
SD	24.40	11.23	11.32	9.00	8.53
SR	0.39	0.61	0.69	0.75	0.84
TO	0.56	1.72	0.68	3.18	1.39
GE	1.00	4.85	4.79	3.66	4.00
			Marko-130-30 Portfolio		
AV	9.51	8.01	7.94	6.24	6.43
SD	24.40	13.10	13.14	9.36	9.31
SR	0.39	0.61	0.60	0.67	0.69
TO	0.56	0.88	0.66	1.77	1.34
GE	1.00	1.60	1.60	1.60	1.60

- Although the relationship is not strictly monotonic, increasing λ tends to increase SD. This makes sense, since the objective is to minimize the variance of the portfolio return subject to a penalty on transaction costs; and the higher the penalty, the less leeway one has in achieving the objective of minimizing the variance.
- The previous two relations imply, on balance, an inverse-U-shaped (and thus concave) relation between λ and SR: As λ increases, typically SR first rises and then it falls again.

Based on these findings, we select $\lambda = 7.5$ as a common good value.

4.5. Results

4.5.1. Portfolios on their own

Armed with these choices, for any portfolio formulation we now present the results for five strategies: the simple-minded benchmark (EW respectively EW-TQ), NL with penalty parameters $\lambda \in \{0, 7.5\}$, and DCC-NL with penalty parameters $\lambda \in \{0, 7.5\}$. Further, to keep the column labeling in the tables compact, we use the following conventions: NL_λ stands for using NL with penalty parameter λ, whereas $DCC\text{-}NL_\lambda$ stands for using DCC-NL with penalty parameter λ. The results are presented in Tables 4.1 and 4.2.

Based on these tables, the following conclusions can be drawn:

- Six of the eight scenarios considered benefit from penalizing transaction costs, that is, from using $\lambda = 7.5$ as opposed to $\lambda = 0$. The two exceptions are the GMV-130-30 and Marko-130-30 portfolios in conjunction with NL. The intuitive reason is that these two scenarios enjoy relatively low turnover already for $\lambda = 0$ so that penalizing the transactions cost does not help but instead hurts (though not by much).

[14] In this way, b_h varies over time, as it makes sense to use a lower value for b_h in a bear market than in a bull market.

[15] As an example for the first goal, if one wants to evaluate the quality of a covariance matrix estimator $\hat{\Sigma}_h$, the canonical method is as follows: Focus on the GMV portfolio without transaction costs, that is, portfolio formulation (2.9) with $\lambda = 0$ and (2.11); use oos instead of oos-ntc returns, that is, ignore transaction costs; then use SD as the leading performance criterion of interest.

- As expected, the improvements from using $\lambda = 7.5$ are consistent and larger for DCC-NL, since this dynamic estimator of the covariance matrix incurs higher turnover compared to the static estimator NL.
- DeMiguel et al. (2009) show that constraining portfolio norms (such as imposing a gross-exposure constraint) improves performance when one uses the sample covariance matrix as the estimator of Σ_h. But if one uses a sophisticated estimator instead, such as NL or DCC-NL, this is no longer necessarily the case. In particular, for DCC-NL the performance of the raw portfolio is superior to the performance of its 130-30 counterpart in both the GMV and the Marko case. On the other hand, for NL the story is mixed: the raw portfolio performs worse in the GMV case but better in the Marko case.
  Arguably, the raw portfolios are not suitable for mainstream players in the industry, such as mutual fund managers, because of their high (average) gross exposure, but they might still be suitable for specialized players who run big leverage ratios, such as hedge-fund managers.
- For all four portfolio formulations, the winner is always the DCC-NL method with $\lambda = 7.5$. In particular, for the raw GMV and Marko portfolios, this method more than doubles the SR of the corresponding simple-minded benchmarks EW and EW-TQ. But even for the constrained GMV-130-30 and Marko-130-30 portfolios, the SR improves by at least 75% relative to the corresponding simple benchmarks.
  DeMiguel et al. (2009) study various portfolio strategies based on estimated input parameters and claim that "none is consistently better than the EW rule [net of transaction costs]". Our results are in clear disagreement with this claim.[16]
- For any given portfolio formulation, starting from the "simplest" portfolio — using the static NL estimator of the covariance matrix and do not account for transaction costs at the portfolio-selection stage ($\lambda = 0$), that is, the NL_0 portfolio — one can think of two separate ways to improve oos portfolio performance: (i) use a dynamic estimator of the covariance matrix (DCC-NL) and (ii) account for transaction costs at the portfolio-selection stage ($\lambda > 0$). In terms of marginal effects, employing (i) alone (DCC-NL_0) is more beneficial than employing (ii) alone ($\text{NL}_{7.5}$). Therefore, even though problem (2.9)–(2.12) is of *static* nature, which we then solve repeatedly across the $H = 275$ portfolio selection dates, it clearly pays off to use a more accurate estimator of the (conditional) covariance matrix, which is the *dynamic* DCC-NL estimator. In terms of overall effect, employing both (i) and (ii) together is the most beneficial approach in the end ($\text{DCC-NL}_{7.5}$).

Remark 4.1 (*Statistical significance*)**.** It might be of interest to study whether the DCC-NL method with $\lambda = 7.5$ delivers a Sharpe ratio that is significantly higher than the DCC-NL method with $\lambda = 0$. To this end, for each of the four portfolio formulations considered, we compute a two-sided p-value for the equality of Sharpe ratios using the prewhitened HAC method of Ledoit and Wolf (2008, Section 3.1). The resulting p-values — in the order of the formulations GMV, GMV-130-30, Marko, and Marko-130-30 — are 0.003, 0.001, 0.054, and 0.596, respectively. Hence, for two of the four formulations we find outperformance at the 5% significance level, and for three of the four formulations we find outperformance at the 10% significance level. □

Remark 4.2 (*Choosing λ for other strategies*)**.** Intuitively, there are many variables that may influence a good choice of the transaction-cost penalty parameter λ in practice, such as the size of the investment universe, the magnitude of (typical) transaction costs, the number and nature of constraints in place, and the portfolio-updating frequency.[17] If, for a given scenario, one considers the Sharpe ratio as a function of λ, then the good news based on our (necessarily) limited results is that the function is rather flat near its peak so that the exact choice of λ is not overly critical, as long as one gets the "neighborhood" right. Therefore, for a given scenario, a backtest over a reasonably long window of past data (say ten to fifteen years of data) should provide enough guidance to choose a good value of λ in practice. As an alternative to the Sharpe ratio criterion, a portfolio manager can also work backwards from an average turnover rate that he/she is comfortable with, and then reverse-engineer through a backtest the value of λ that delivers it.

If one is concerned that the "optimal' value of λ might change over time, a rolling-window approach based on a relatively short evaluation window, of say y years, can be taken. In this case one needs a (total) window going back $5 + y$ years; this is because at the beginning of the evaluation period one needs five years of past data to estimate the covariance matrix (ideally using DCC-NL). Next one runs a backtest as above over the evaluation period of y years to determine the "optimal" λ. This value of λ is then used for a period of m months. Then one rolls forward m months and repeats the procedure. In machine learning lingo, the period of five years (to estimate the covariance matrix and perhaps also the vector of expected returns) corresponds to the *training set*, the evaluation period of y years to determine the "optimal" λ corresponds to the *validation* set, and the period of m months to compute (pseudo) oos returns corresponds to the *test set*. One can even run a "hyper backtest" to determine "optimal" values of y and m. In doing such an exercise one clearly moves from the academic realm to the practitioner realm; hence, we shall abstain from such a "hyper backtest" in this paper. □

4.5.2. Distance between portfolios

In this section, we study the average distance of a portfolio for a given formulation (GMV, GMV-130-30, Marko, Marko-130-30) from two reference portfolios as a function of the penalty parameter λ. The two reference portfolios are (i) the corresponding portfolio with $\lambda = 0$ and (ii) the equally-weighted (EW) portfolio. We now describe how to compute the average distance for a given combination of portfolio formulation can covariance-matrix estimator. For a portfolio-selection date h and a value of the penalty parameter λ, denote the portfolio weights by $w_{\lambda,h,i}$ and the reference-portfolio weights by $w_{\text{ref},h,i}$, for $i = 1, \ldots, N$. We then compute the distance[18] between the two portfolios, defined as the sum of the absolute differences of the respective portfolio weights, by

$$|d_{\lambda,h}| := \|w_{\lambda,h} - w_{\text{ref},h}\|_1 = \sum_{i=1}^{N} |w_{\lambda,h,i,} - w_{\text{ref},h,i}| \tag{4.1}$$

and average these distances over the 274 portfolio-selection dates $h = 2, \ldots, 275$ to obtain the average-distance measure

$$\overline{|d_\lambda|} := \frac{1}{274} \sum_{h=2}^{275} |d_{\lambda,h}| \, . \tag{4.2}$$

(We exclude $h = 1$ in this average because, by definition, the portfolio weights $w_{1,\lambda}$ are equal across λ.) Finally, we plot the measures $\overline{|d_\lambda|}$ against $\lambda \in \{0, 2.5, \ldots, 20, 50\}$. There are a total of 16 such plots:

[16] Part of the reason is that they consider an unduly large transaction cost of 50 bps (constant across stocks); such a cost may have been realistic decades ago but it is no longer anywhere near the costs investors face today. Another part of the reason is the particular list of strategies the authors consider, for reasons not clear, leaves out some promising contenders.

[17] *Ceteris paribus*, the more frequently a portfolio is updated, such as daily vs. monthly updating, the larger a good value of λ needs to be.

[18] One could also interpret this distance as a *turnover* if one were to turn over the portfolio defined by the weigh vector $w_{\lambda,h}$ to the reference portfolio defined by the weight vector $w_{\text{ref},h}$, or vice versa.

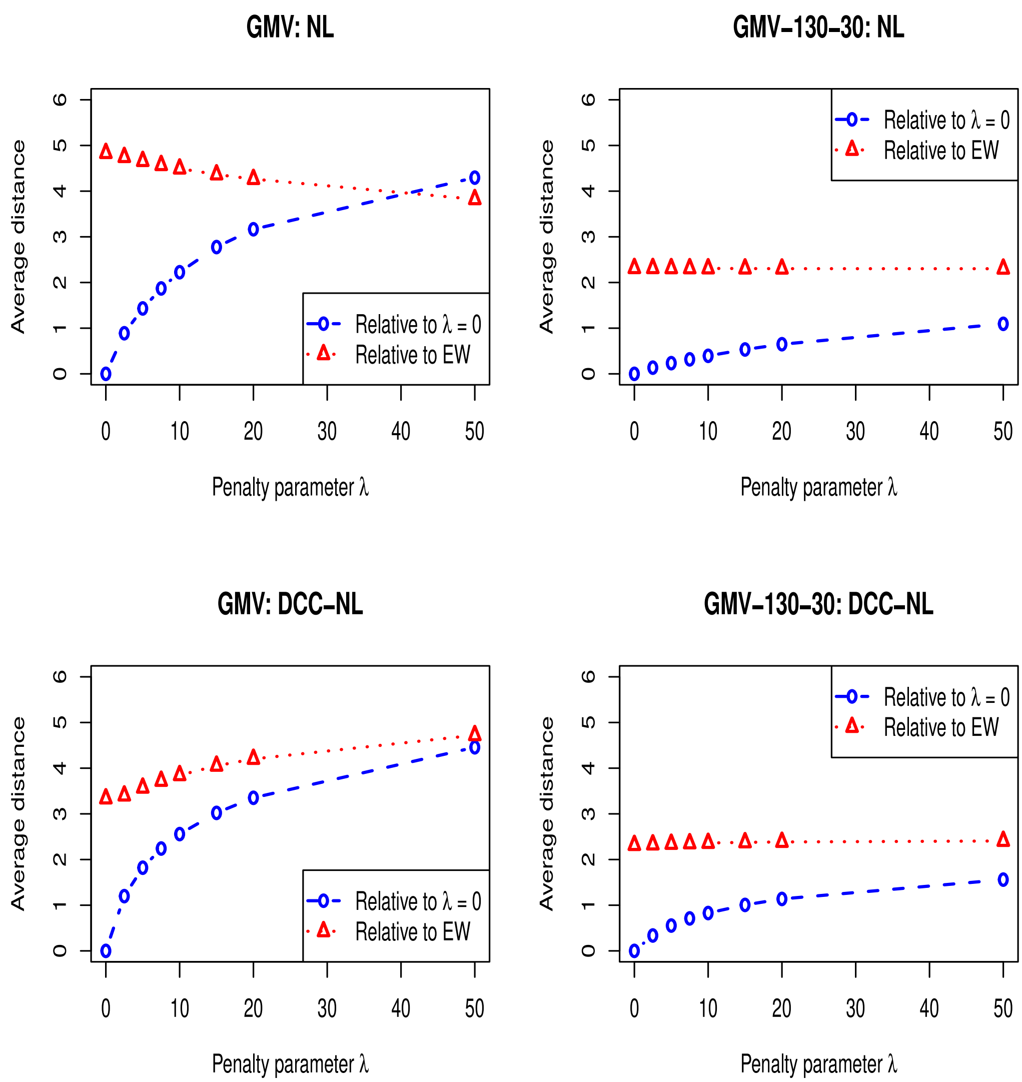

Fig. 4.1. Average-distance measure (4.2) as a function of the penalty parameter λ relative to two reference portfolios. These graphs are for the four GMV-type portfolio formulations.

- four portfolio formulations: GMV, GMV-130-30, Marko, and Marko-130-30; and
- two covariance-matrix estimators: NL and DCC-NL; and
- two reference portfolios: $\lambda = 0$ and EW.

Fig. 4.1 displays the eight plots pertaining to GMV-type formulations whereas Fig. 4.2 displays the eight plots pertaining to Marko-type formulations. The results can be summarized as follows:

- In any scenario, meaning for any combination of portfolio formulation and covariance-matrix estimator, the average-distance measure $\overline{|d_\lambda|}$ relative to the reference portfolio $\lambda = 0$ increases in λ. This makes sense that, since the larger the value of λ is, the more is the portfolio forced to stay close to the holding portfolio, as opposed to getting close to the reference portfolio $\lambda = 0$.
- On the other hand, increasing the value of λ does not mean that the portfolio gets closer to the reference portfolio EW. In many scenarios, in particular all those with a 130-30 constraint, the average-distance measure $\overline{|d_\lambda|}$ relative to the EW portfolio is almost constant across λ. In the scenarios without 130-30 constraint, the average-distance measure $\overline{|d_\lambda|}$ relative to the EW portfolio decreases in λ in one scenario, is nearly constant across λ in one scenario, and actually increases in λ in two scenarios. Hence, on balance, increasing λ does not mean getting closer to EW; instead, it means staying closer to the holding portfolio.
- As can be clearly seen, average distances relative to the reference portfolio $\lambda = 0$ are generally much lower for scenarios with a 130-30 constraint. This makes sense, since by the triangle inequality the distance (4.1), and therefore also the average-distance measure (4.2), is bounded above by 3.2 in such a case. On the other hand, without a 130-30 constraint, there is no (guaranteed) upper bound. To look at it from a slightly different angle, portfolio formulations with larger typical gross exposures also lead to larger typical average distances between a portfolio for a given λ and the reference portfolio $\lambda = 0$.

5. Conclusion

In this paper, we have proposed a method to account for transaction costs at the portfolio-selection stage in the context of Markowitz portfolio selection. Doing so often increases the Sharpe ratio net of

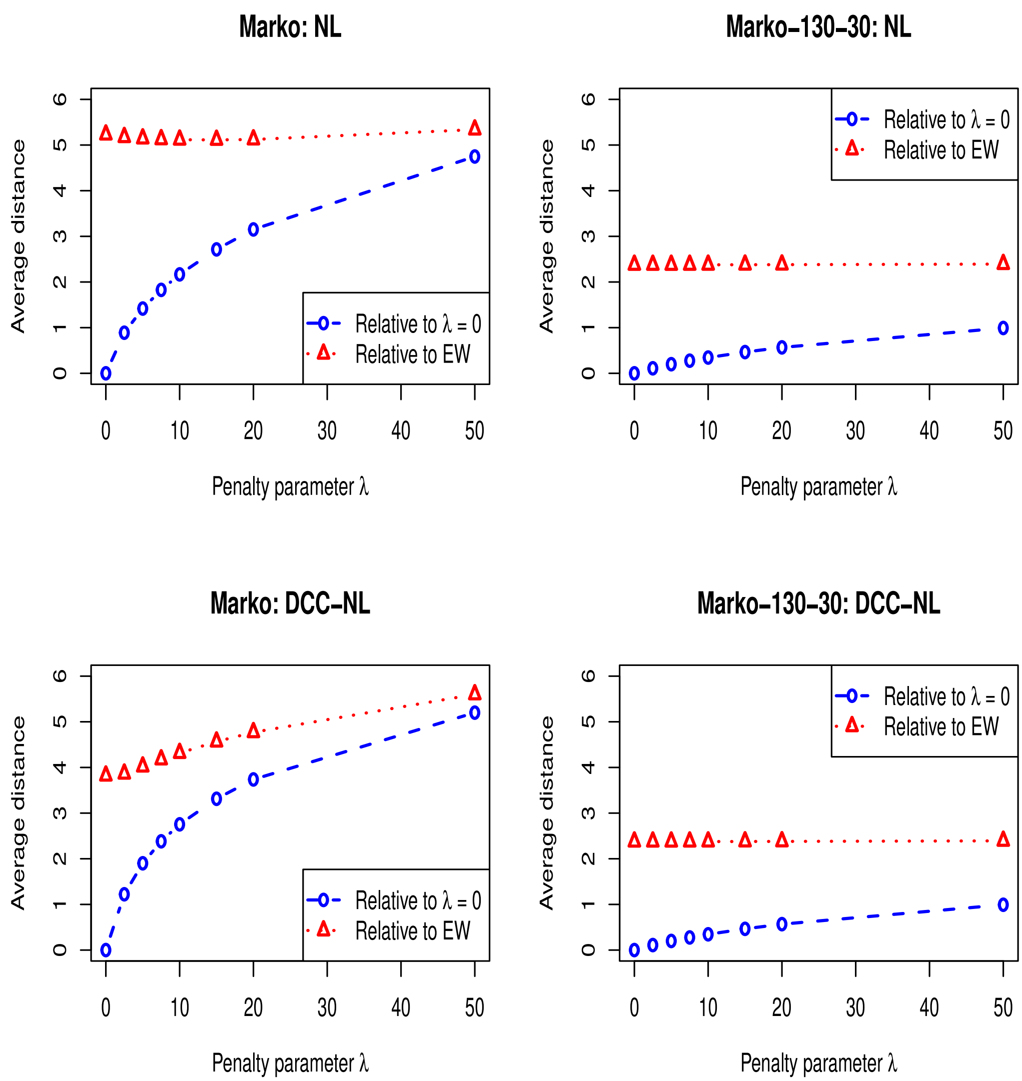

Fig. 4.2. Average-distance measure (4.2) as a function of the penalty parameter λ relative to two reference portfolios. These graphs are for the four Marko-type portfolio formulations.

transaction costs compared to the *status quo* in the literature which amounts to ignoring transaction costs at the portfolio-selection stage and simply paying them after the fact. Importantly, and also deviating from the *status quo* in the literature, we incorporate the fact that transaction costs are stock-specific instead of taking the easy route of assuming them to be equal across stocks. The benefit of our proposal is especially pronounced when dynamic estimators of the covariance matrix are used as inputs in the Markowitz formulations. The reason is that dynamic estimators, although being more accurate, typically generate higher turnover compared to static estimators and, therefore, investment strategies based upon them suffer more from transaction costs. We hope that our proposal will be of future use to anyone interested in devising Markowitz portfolio-selection strategies that have desirable performance net of transaction costs.

Although the focus of this paper has been on the covariance matrix, an obvious avenue for future research is to use the technology introduced here to further tame the factor zoo (Feng et al., 2020) by weeding out those alphas that are over-dependent on high-transaction-cost stocks and turn over too much to pay for themselves — but only those. In this context, ignoring transaction costs is too lenient, whereas charging a constant transaction cost to every stock is unrealistic, so our method might prove to be just the right "scalpel".

Declaration of competing interest

We have no conflicts of interest to declare.

Appendix A. Non-Markowitz paradigms: Two examples

A.1. Post-processing the standard Markowitz solution

As a representative example of this paradigm, Frahm and Memmel (2010) estimate the unconstrained GMV portfolio, which is the portfolio given by (2.1) and (2.3), dropping the constraint (2.2). This problem has the following analytical solution:

$$w^{\text{GMV}} = \frac{\mathbb{1}' \Sigma^{-1}}{\mathbb{1}' \Sigma^{-1} \mathbb{1}} .$$

Denote the sample covariance matrix by S. Then the corresponding plug-in solution is given by

$$\hat{w} = \frac{\mathbb{1}' S^{-1}}{\mathbb{1}' S^{-1} \mathbb{1}} .$$

(Here we tacitly assume that the number of observations exceeds the number of stocks, since otherwise S is not invertible.) Unless the number of stocks is very small relative to the number of observations, this solution is know to have poor oos performance. Frahm and Memmel (2010) propose to modify it as follows:

$$\hat{w}^{\text{mod}} := \alpha w^{\text{EW}} + (1-\alpha)\hat{w} \quad \text{where} \quad w^{\text{EW}} := (1/N, \ldots, 1/N)' .$$

Therefore, the proposed modification is a convex combination of the equally-weighted (EW) portfolio and the sample-based solution $\hat{w}$; another way to look at it is that the sample-based solution $\hat{w}$ is linearly shrunk towards the EW portfolio. In practice, one has to work out a suitable choice of the shrinkage intensity $\alpha \in [0,1]$ and Frahm and Memmel (2010) offer a corresponding solution.

We point that (i) this method cannot handle the case $N > T$ because it requires the sample covariance matrix to be of full rank; (ii) it is not clear how to account for transaction costs at the portfolio-selection stage and (iii) it is not clear how to incorporate additional constraints such as (as the leading case) a gross-exposure constraint.

A.2. Specifying a different optimization program

As a representative example of this paradigm, consider the following problem formulation:

$$\max_{w \in \mathbb{R}^N} w'\mu \tag{A.1}$$

$$\text{subject to} \quad w'\Sigma w \leq \sigma^2 , \tag{A.2}$$

where σ^2 is an upper bound on the portfolio variance. This problem has the following analytical solution:

$$w^{\text{SR}} := \frac{\sigma}{\sqrt{\theta}} \Sigma^{-1}\mu \quad \text{where} \quad \theta := \mu'\Sigma^{-1}\mu . \tag{A.3}$$

(When the returns are expressed in excess of the risk-free rate, (A.3) is also known as the maximum Sharpe ratio (MSR) portfolio.) The standard feasible solution would be to plug in the sample mean vector for μ and the sample covariance matrix for Σ in (A.3); but, as discussed before, the resulting portfolio performs poorly out of sample. Instead the proposal of Ao et al. (2019) works as follows. First, the problem (A.1)–(A.2) can be re-expressed as an unconstrained population regression problem. Second, the population regression problem can be approximated by a sample analog (based on past observations). Third, the regression problem involves the unknown parameter θ which (under normality) can be estimated in an unbiased fashion as

$$\hat{\theta} := \frac{(T-N-2)\hat{\theta}_s - N}{T} , \tag{A.4}$$

where T denotes the number of observations, N denotes the number of stocks, and $\hat{\theta}_s$ denotes the sample analog of θ. Fourth, the sample analog of the regression problem is estimated via the LASSO instead of OLS. A crucial feature is that no sophisticated estimators of μ and Σ are needed.

We point that (i) this method cannot handle the case $N > T$ because then formula (A.4) results in a negative estimator $\hat{\theta}$; (ii) it is not clear how to account for transaction costs at the portfolio-selection stage and (iii) it is not clear how to incorporate additional constraints such as (as the leading case) a gross-exposure constraint.

Appendix B. Additional tables

See Tables B.1–B.4.

Table B.1
Annualized performance measures (in percent) for various estimators of the GMV portfolio. AV stands for average; SD stands for standard deviation; and SR stands for Sharpe ratio. All measures are based on 5,775 daily out-of-sample returns net of transaction costs and in excess of the risk-free rate. Furthermore, TO stands for average monthly turnover and GE stands for average gross exposure.

λ	0	2.5	5.0	7.5	10	15	20	50
				GMV Portfolio: NL				
AV	6.75	7.33	7.60	7.71	7.77	7.70	7.93	8.38
SD	9.89	9.87	9.86	9.86	9.88	9.94	10.01	10.46
SR	0.68	0.74	0.77	0.78	0.79	0.77	0.79	0.80
TO	1.30	0.66	0.47	0.38	0.32	0.25	0.21	0.12
GE	4.41	4.34	4.29	4.24	4.19	4.11	4.04	3.73
				GMV Portfolio: DCC-NL				
AV	7.63	7.74	7.70	7.69	7.59	7.69	7.84	7.73
SD	8.14	7.52	7.27	7.20	7.19	7.24	7.32	7.92
SR	0.94	1.03	1.06	1.07	1.06	1.06	1.07	0.98
TO	2.84	1.69	1.29	1.08	0.95	0.79	0.69	0.47
GE	3.14	3.20	3.36	3.51	3.62	3.81	3.95	4.39

Table B.2
Annualized performance measures (in percent) for various estimators of the GMV-130-30 portfolio. AV stands for average; SD stands for standard deviation; and SR stands for Sharpe ratio. All measures are based on 5,775 daily out-of-sample returns net of transaction costs and in excess of the risk-free rate. Furthermore, TO stands for average monthly turnover and GE stands for average gross exposure.

λ	0	2.5	5.0	7.5	10	15	20	50
				GMV-130-30 Portfolio: NL				
AV	8.84	8.77	8.70	8.61	8.47	8.25	8.11	8.21
SD	10.51	10.54	10.57	10.59	10.63	10.70	10.76	11.01
SR	0.84	0.83	0.82	0.81	0.80	0.77	0.75	0.75
TO	0.45	0.33	0.27	0.24	0.21	0.18	0.16	0.10
GE	1.60	1.60	1.60	1.60	1.60	1.60	1.60	1.60
				GMV-130-30 Portfolio: DCC-NL				
AV	6.82	7.06	7.14	7.06	7.07	7.12	7.07	7.36
SD	7.80	7.64	7.55	7.46	7.40	7.35	7.38	7.73
SR	0.87	0.92	0.95	0.95	0.96	0.97	0.96	0.95
TO	1.76	1.43	1.24	1.11	1.02	0.89	0.80	0.54
GE	1.60	1.60	1.60	1.60	1.60	1.60	1.60	1.60

Table B.3
Annualized performance measures (in percent) for various estimators of the Marko portfolio. AV stands for average; SD stands for standard deviation; and SR stands for Sharpe ratio. All measures are based on 5,775 daily out-of-sample returns net of transaction costs and in excess of the risk-free rate. Furthermore, TO stands for average monthly turnover and GE stands for average gross exposure.

λ	0	2.5	5.0	7.5	10	15	20	50
				Marko Portfolio: NL				
AV	6.88	7.66	7.91	7.82	7.72	7.41	7.24	6.91
SD	11.23	11.23	11.26	11.32	11.39	11.56	11.71	12.68
SR	0.61	0.68	0.70	0.69	0.68	0.64	0.62	0.54
TO	1.72	1.05	0.81	0.68	0.60	0.50	0.44	0.31
GE	4.85	4.81	4.79	4.79	4.79	4.81	4.83	5.06
				Marko Portfolio: DCC-NL				
AV	6.75	7.27	7.21	7.16	6.97	6.97	7.07	6.87
SD	9.00	8.60	8.50	8.53	8.58	8.78	8.99	9.90
SR	0.75	0.84	0.85	0.84	0.81	0.79	0.79	0.69
TO	3.18	2.05	1.62	1.39	1.24	1.06	0.95	0.69
GE	3.66	3.70	3.85	4.00	4.14	4.37	4.57	5.36

Table B.4
Annualized performance measures (in percent) for various estimators of the Marko-130-30 portfolio. AV stands for average; SD stands for standard deviation; and SR stands for Sharpe ratio. All measures are based on 5,775 daily out-of-sample returns net of transaction costs and in excess of the risk-free rate. TO stands for average turnover and GE stands for average gross exposure.

λ	0	2.5	5.0	7.5	10	15	20	50
				Marko-130-30 Portfolio: NL				
AV	8.01	8.03	7.99	7.94	7.86	7.77	7.64	7.32
SD	13.10	13.10	13.12	13.14	13.17	13.24	13.31	13.80
SR	0.61	0.61	0.61	0.60	0.60	0.59	0.57	0.53
TO	0.88	0.78	0.71	0.66	0.61	0.55	0.50	0.36
GE	1.60	1.60	1.60	1.60	1.60	1.60	1.60	1.60
				Marko-130-30 Portfolio: DCC-NL				
AV	6.24	6.46	6.52	6.43	6.46	6.44	6.41	6.49
SD	9.36	9.30	9.28	9.31	9.38	9.46	9.62	10.79
SR	0.67	0.69	0.70	0.69	0.69	0.68	0.67	0.60
TO	1.77	1.57	1.44	1.34	1.26	1.15	1.06	0.80
GE	1.60	1.60	1.60	1.60	1.60	1.60	1.60	1.60

References

Abdi, F., & Ranaldo, A. (2017). A simple estimation of bid-ask spreads from daily close, high, and low prices. *The Review of Financial Studies*, *30*(12), 4437–4480.

Ao, M., Li, Y., & Zheng, X. (2019). Approaching mean-variance efficiency for large portfolios. *The Review of Financial Studies*, *32*(7), 2890–2919.

Ardia, D., Guidotti, E., & Kroencke, T. A. (2024). Efficient estimation of bid–ask spreads from open, high, low, and close prices. *Journal of Financial Economics*, *161*, Article 103916.

Brandt, M. W., Santa-Clara, P., & Valkanov, R. (2009). Parametric portfolio policies: Exploiting characteristics in the cross-section of equity returns. *The Review of Financial Studies*, *22*(9), 3411–3447.

Briere, M., Lehalle, C.-A., Nefedova, T., & Raboun, A. (2020). Modeling transaction costs when trades may be crowded: A Bayesian network using partially observable orders imbalance. In E. Jurczenko (Ed.), *Machine Learning for Asset Management: New developments and Financial Applications* (pp. 387–430). Hoboken: John Wiley.

Chen, Z., Ding, Y., Li, Y., & Zheng, X. (2024). *Learning the stochastic discount factor*: *Working paper*, Hong Kong University of Science and Technology.

DeMiguel, V., Garlappi, L., Nogales, F. J., & Uppal, R. (2009). A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms. *Management Science*, *55*(5), 798–812.

DeMiguel, V., Garlappi, L., & Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the $1/N$ portfolio strategy? *The Review of Financial Studies*, *22*, 1915–1953.

DeMiguel, V., Martin-Utrera, A., Nogales, F. J., & Uppal, R. (2020). A transaction-cost perspective on the multitude of firm characteristics. *The Review of Financial Studies*, *33*(5), 2180–2222.

Dixon, M. F., Halperin, I., & Bilokon, P. (2020). *Machine Learning in Finance*: *vol. 1170*, New York: Springer International Publishing.

Engle, R. F., Ledoit, O., & Wolf, M. (2019). Large dynamic covariance matrices. *Journal of Business & Economic Statistics*, *37*(2), 363–375.

Feng, G., Giglio, S., & Xiu, D. (2020). Taming the factor zoo: A test of new factors. *The Journal of Finance*, *75*(3), 1327–1370.

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

Gârleanu, N., & Pedersen, L. H. (2013). Dynamic trading with predictable returns and transaction costs. *The Journal of Finance*, *68*(6), 2309–2340.

Garman, M. B., & Klass, M. J. (1980). On the estimation of security price volatilities from historical data. *Journal of Business*, *53*(1), 67–78.

Green, J., Hand, J. R. M., & Zhang, X. F. (2013). The supraview of return predictive signals. *Review of Accounting Studies*, *18*, 692–730.

Harvey, C. R., Liu, Y., & Zhu, H. (2016). ... and the cross-section of expected returns. *The Review of Financial Studies*, *29*(1), 5–68.

Hasbrouck, J. (2009). Trading costs and returns for US equities: Estimating effective costs from daily data. *The Journal of Finance*, *64*(3), 1445–1477.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance*, *48*(1), 65–91.

Kan, R., Wang, X., & Zhou, G. (2022). Optimal portfolio choice with estimation risk: No risk-free asset case. *Management Science*, *68*(3), 2047–2068.

Kan, R., & Zhou, G. (2007). Optimal portfolio choice with parameter uncertainty. *Journal of Financial and Quantitative Analysis*, *42*(3), 621–656.

Kelly, B., & Xiu, D. (2023). Financial machine learning. *Foundations and Trends® in Finance*, *13*(3–4), 205–363.

Kolm, P. N., Tütüncü, R., & Fabozzi, F. J. (2014). 60 years of portfolio optimization: Practical challenges and current trends. *European Journal of Operational Research*, *234*(2), 356–371.

Ledoit, O., & Wolf, M. (2008). Robust performance hypothesis testing with the Sharpe ratio. *Journal of Empirical Finance*, *15*, 850–859.

Ledoit, O., & Wolf, M. (2011). Robust performance hypothesis testing with the variance. *Wilmott Magazine*, *September*, 86–89.

Ledoit, O., & Wolf, M. (2022a). The power of (non-)linear shrinking: A review and guide to covariance matrix estimation. *Journal of Financial Econometrics*, *20*(1), 187–218.

Ledoit, O., & Wolf, M. (2022b). Quadratic shrinkage for large covariance matrices. *Bernoulli*, *28*(3), 1519–1547.

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

López de Prado, M. M. (2016). Building diversified portfolios that outperform out-of-sample. *Journal of Portfolio Management*, *42*(4), 59–69.

López de Prado, M. M. (2020). *Machine Learning for Asset Aanagers*. Cambridge: Cambridge University Press.

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

Michaud, R. (1989). The Markowitz optimization enigma: Is optimized optimal? *Financial Analysts Journal*, *45*, 31–42.

Neely, C. J., Rapach, D. E., Tu, J., & Zhou, G. (2014). Forecasting the equity risk premium: the role of technical indicators. *Management Science*, *60*(7), 1772–1791.

Novy-Marx, R., & Velikov, M. (2016). A taxonomy of anomalies and their trading costs. *The Review of Financial Studies*, *29*(1), 104–147.

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