Turun yliopisto
University of Turku

ABSTRACT

- [x] Master´s thesis
- [ ] Licentiate´s thesis
- [ ] Doctor´s thesis

Subject	Accounting and Finance	Date	26.1.2022
Author(s)	Ville Kaukonen	Student number	418955
		Number of pages	220
Title	Survey of quantitative investment strategies in the Russian stock market: Special interest in tactical asset allocation		
Supervisor(s)	Prof. Mika Vaihekoski, Ph.D. Henri Teittinen and M.Sc. Mika Jakovaara		

Abstract

Russia’s financial markets have been an uncharted area when it comes to exploring the performance of investment strategies based on modern portfolio theory. In this thesis, we focus on the country’s stock market and study whether profitable investments can be made while at the same time taking uncertainties, risks, and dependencies into account. We also pay particular interest in tactical asset allocation. The benefit of this approach is that we can utilize time series forecasting methods to produce trading signals in addition to optimization methods.

We use two datasets in our empirical applications. The first one consists of nine sectoral indices covering the period from 2008 to 2017, and the other includes altogether 42 stocks listed on the Moscow Exchange covering the years 2011 – 2017. The strategies considered have been divided into five sections. In the first part, we study classical and robust mean-risk portfolios and the modeling of transaction costs. We find that the expected return should be maximized per unit expected shortfall while simultaneously requiring that each asset contributes equally to the portfolio’s tail risk. Secondly, we show that using robust covariance estimators can improve the risk-adjusted returns of minimum variance portfolios. Thirdly, we note that robust optimization techniques are best suited for conservative investors due to the low volatility allocations they produce.

In the second part, we employ statistical factor models to estimate higher-order comoments and demonstrate the benefit of the proposed method in constructing risk-optimal and expected utility-maximizing portfolios. In the third part, we utilize the Almgren–Chriss framework and sort the expected returns according to the assumed momentum anomaly. We discover that this method produces stable allocations performing exceptionally well in the market upturn. In the fourth part, we show that forecasts produced by VECM and GARCH models can be used profitably in optimizations based on the Black–Litterman, copula opinion pooling, and entropy pooling models. In the final part, we develop a wealth protection strategy capable of timing market changes thanks to the return predictions based on an ARIMA model.

Therefore, it can be stated that it has been possible to make safe and profitable investments in the Russian stock market even when reasonable transaction costs have been taken into account. We also argue that market inefficiencies could have been exploited by structuring statistical arbitrage and other tactical asset allocation-related strategies.

Key words	investment strategies, portfolio optimization, Russia, stock market, risk management
Further information	

Turun yliopisto
University of Turku

TIIVISTELMÄ

- [x] Pro gradu -tutkielma
- [ ] Lisensiaatintutkielma
- [ ] Väitöskirja

Oppiaine	Laskentatoimi ja rahoitus	Päivämäärä	26.1.2022
Tekijä(t)	Ville Kaukonen	Matrikkelinumero	418955
		Sivumäärä	220
Otsikko	Selvitys kvantitatiivisista sijoitusstrategioista Venäjän osakemarkkinoilla: Erityinen mielenkiinto taktiseen varojen kohdentamiseen		
Ohjaaja(t)	Prof. Mika Vaihekoski, KTT Henri Teittinen ja KTM Mika Jakovaara		

Tiivistelmä

Venäjän rahoitusmarkkinat ovat olleet kartoittamatonta aluetta tutkittaessa moderniin portfolioteoriaan pohjautuvien sijoitusstrategioiden käyttäytymistä. Tässä tutkielmassa keskitymme maan osakemarkkinoihin ja tarkastelemme, voidaanko taloudellisesti kannattavia sijoituksia tehdä otettaessa samalla huomioon epävarmuudet, riskit ja riippuvuudet. Kiinnitämme erityistä huomiota myös taktiseen varojen kohdentamiseen. Tämän lähestymistavan etuna on, että optimointimenetelmien lisäksi voimme hyödyntää aikasarjaennustamisen menetelmiä kaupankäyntisignaalien tuottamiseksi.

Empiirisissä sovelluksissa käytämme kahta data-aineistoa. Ensimmäinen koostuu yhdeksästä teollisuusindeksistä kattaen ajanjakson 2008–2017, ja toinen sisältää 42 Moskovan pörssiin listattua osaketta kattaen vuodet 2011–2017. Tarkasteltavat strategiat on puolestaan jaoteltu viiteen osioon. Ensimmäisessä osassa tarkastelemme klassisia ja robusteja riski-tuotto -portfolioita sekä kaupankäyntikustannusten mallintamista. Havaitsemme, että odotettua tuottoa on syytä maksimoida suhteessa odotettuun vajeeseen edellyttäen samalla, että jokainen osake lisää sijoitussalkun häntäriskiä yhtä suurella osuudella. Toiseksi osoitamme, että minimivarianssiportfolioiden riskikorjattuja tuottoja voidaan parantaa robusteilla kovarianssiestimaattoreilla. Kolmanneksi toteamme robustien optimointitekniikoiden soveltuvan parhaiten konservatiivisille sijoittajille niiden tuottamien matalan volatiliteetin allokaatioiden ansiosta.

Toisessa osassa hyödynnämme tilastollisia faktorimalleja korkeampien yhteismomenttien estimoinnissa ja havainnollistamme ehdotetun metodin hyödyllisyyttä riskioptimaalisten sekä odotettua hyötyä maksimoivien salkkujen rakentamisessa. Kolmannessa osassa käytämme Almgren–Chrissin viitekehystä ja asetamme odotetut tuotot suuruusjärjestykseen oletetun momentum-anomalian mukaisesti. Havaitsemme, että menetelmä tuottaa vakaita allokaatioita menestyen erityisen hyvin noususuhdanteessa. Neljännessä osassa osoitamme, että VECM- että GARCH-mallien tuottamia ennusteita voidaan hyödyntää kannattavasti niin Black–Littermanin malliin kuin kopulanäkemysten ja entropian poolaukseenkin perustuvissa optimoinneissa. Viimeisessä osassa laadimme varallisuuden suojausstrategian, joka kykenee ajoittamaan markkinoiden muutoksia ARIMA-malliin perustuvien tuottoennusteiden ansiosta.

Voidaan siis todeta, että Venäjän osakemarkkinoilla on ollut mahdollista tehdä turvallisia ja tuottavia sijoituksia myös silloin kun kohtuulliset kaupankäyntikustannukset on huomioitu. Toiseksi väitämme, että markkinoiden tehottomuutta on voitu hyödyntää suunnittelemalla tilastolliseen arbitraasiin ja muihin taktiseen varojen allokointiin pohjautuvia strategioita.

Asiasanat	sijoitusstrategiat, portfolio-optimointi, Venäjä, osakemarkkinat, riskienhallinta
Muita tietoja	

Turun yliopisto
University of Turku

SURVEY OF QUANTITATIVE INVESTMENT STRATEGIES IN THE RUSSIAN STOCK MARKET

Special interest in tactical asset allocation

Master´s Thesis
in Accounting and Finance

Author:
Ville Kaukonen

Supervisors:
Prof. Mika Vaihekoski
Ph.D. Henri Teittinen
M.Sc. Mika Jakovaara

26.1.2022
Turku

Turun kauppakorkeakoulu • Turku School of Economics

Table of contents

List of figures

List of tables

1 INTRODUCTION

1.1 Motivation and background

This thesis surveys the performance of different asset allocation strategies in the emerging Russian stock market. *Asset allocation* can be defined as a decision on how to allocate the wealth of an investor between the set of risky securities and possible risk-free security. Hence, it is essentially based on the classical theory of portfolio selection by Markowitz (1952) and risk-reward optimization. Moreover, asset allocation is an integral part of designing an investment plan and policy for an investor, which, in turn, is closely related to the strategic goals, general risk profile, and liquidity requirements of an investor. (Fabozzi et al. 2010, 325, 392, 417.)

In order to shed more light on the dynamics of the Russian market, special attention is also paid to *Tactical Asset Allocation* (TAA). According to Lustig (2013, 239–242), tactical asset allocation is the process of deviating from the weights of Strategic Asset Allocation (SAA), which is the portfolio's long-term investment strategy. The underlying objective is to enhance performance by benefiting from relative value and market opportunities and managing and mitigating risks. Thus, in addition to the fact that TAA seeks to exploit short-term fundamental inefficiencies or temporary imbalances in equilibrium values, it can also be seen as a risk management tool because it can efficiently and quickly alter the asset allocation to control the risk exposure (see also Tokat, Wicas & Stockton 2007, 33–35). Similarly, Pfaff (2016, 274) defines TAA as a set of techniques used to shift assets in and out of a portfolio in a systematic manner. In other words, the aim is first to outline how relative and directional forecasts can be obtained and then how portfolio allocations can be derived from a set of quantitatively deduced signals.

Initially, interest in dynamic TAA strategies over static buy-and-hold approaches grew when evidence of the forecastability of stock and bond market returns appeared in the literature (see, e.g., Merton 1981; Breen, Glosten & Jagannathan 1989; Campbell 1987; Fama & French 1988; Fama & French 1989; Weigel 1991; Philips, Rogers & Capaldi 1996). As a result, market and volatility timing, which refers to strategies allowing investors to allocate their wealth using return and volatility forecasts, have been studied extensively in the Western markets (see, e.g., Brocato & Chandy 1994; Kandel & Stambaugh 1996; Barberis 2000; Fleming, Kirby & Ostdiek 2001 & 2003; Johannes, Polson & Stroud 2002; Li & Lam 2002; Marquering & Verbeek 2004; Han 2006; Chou & Liu 2010; Jondeau & Rockinger 2012; Clements & Silvennoinen 2013). For example, Barberis (2000) examines how the evidence of predictability in asset returns affects optimal portfolio choice for investors with long horizons. He finds that even after incorporating parameter uncertainty, there is enough predictability in returns to make investors allocate substantially more to stocks, the longer their horizon. Johannes et al. (2002), in turn, find that a strategy based merely on volatility timing uniformly outperforms market timing strategies, as well as the market return in terms of Sharpe ratios and certainty equivalent gains. According to the authors, market timing strategies perform poorly due to estimation risk, which is the considerable uncertainty present in estimating and forecasting expected returns. Marquering and Verbeek (2004) complement this notion by finding that returns' predictability is larger when volatility is high, which is consistent with the existence of time-varying risk premia, as well as with incomplete learning after a significant shock to the economy. Finally, Jondeau and Rockinger (2012) show that distribution timing, defined as the ability to use forecasts for moments up to the fourth one, generates significant incremental economic value.

It is worth emphasizing that there are several articles about wealth management that use the term tactical asset allocation but in slightly different contexts (see, e.g., Phillips & Lee 1989; MacBeth & Emanuel 1993; Lee 1998; Lee 2000ab; Casas 2001; Qian 2003; Anson 2004; Herold & Maurer 2004; Rey 2004ab; Roberge & Le Moigne 2005; Lewis, Okunev & White 2007; Dichtl & Drobetz 2009; Wang & Kochard 2012; Chong & Phillips 2014; Chakraborty, Grant, and Trahan (2017); Yang et al. 2018). Even though this causes some variation to the definition of the term, it most often involves testing such investment strategies where financial or economic analysis is used to predict or anticipate future performance and assign relative short- to near-term asset-class weightings accordingly.

For example, Yang et al. (2018) find that no TAA strategies on technical trading rules can outperform the buy-and-hold benchmark. Wang and Kochard (2012), in turn, use a Z-score approach to investigate whether momentum and value effects can be observed across asset classes at the index level in a tactical asset allocation framework. They show that it is possible to improve the risk-adjusted returns on a given strategic asset allocation by tactically adjusting the relative weights of asset classes based on their perceived momentum and value attractiveness.

Furthermore, Herold and Maurer (2004) focus on sector rotation strategies and employ instrumental variables (such as dividend yields and terms spreads) to predict expected returns. They find that incorporating estimation risk (i.e., parameter uncertainty) reduces turnover and improves the risk-adjusted performance of dynamic and active asset allocation strategies. The most promising strategies also seem to be based on Bayesian statistics. This is perhaps one of the reasons why the Black–Litterman model (see Black & Litterman 1991ab; 1992) has become the most widely known and almost synonymous with TAA-driven portfolio allocations. Despite this, tactical asset allocation can also be combined with, e.g., risk-overlay and high-watermark strategies (Pfaff 2016, 274).

Indeed, in contrast to the traditional approaches where TAA is based upon forecasting asset returns (e.g., Casas 2001), Lewis et al. (2007) present a way to include a Value-at-

Risk (VaR) constraint to tactical asset allocation. The dynamic VaR TAA strategy they propose provides investors with prescribed tactical tilts in asset allocation that align with their risk aversion level. Hence, it can be utilized to control the risk and expected losses of any balanced product. Moreover, since this strategy does not require estimating an investor's utility function, it offers an attractive alternative to the approaches used by Lee (2000ab) and Rey (2004ab).

Of course, the paper by Lewis et al. (2007) can also be considered to resemble closely such rules-based strategies that do not rely on forecasts and are usually designed to produce absolute returns, i.e., either positive returns or total returns above a pre-specified minimum return (Dichtl & Drobetz 2009, 236). In brief, rules-based strategies, which seek to dynamically manage portfolio risk in order to protect the portfolio value from falling below a predefined level or floor, encompass shortfall risk-based approaches (e.g., Harlow 1991; Herold et al. 2007), as well as various portfolio insurance techniques (e.g., Rubinstein & Leland 1981; Choie & Seff 1989; Zimmermann 1996; Annaert, Van Osselaer & Verstraete 2009; see also Dichtl & Drobetz 2011 for a comparison of option- and constant proportion-based portfolio protection strategies).

As stated earlier, our focus is on the Russian equities, meaning that we omit the considerations of other asset classes, such as fixed income (bond portfolio management strategies are discussed, e.g., in Fabozzi 2013). This delimitation can be advocated by the findings of An, Ang, and Collin-Dufresne (2015): They calculate optimal tactical asset allocation policies over equities as well as bonds when both asset returns are predictable. By changing how often one resets the weights, the authors estimate the benefits and costs of different frequencies of TAA decisions. They conclude that tactical tilts taking advantage of predictable stock returns produce approximately twice as much value as those market-timing bond returns.

Arguably, the Russian stock market provides a problematic but novel and uncharted arena for examining the performance of investment strategies: First of all, in Russia, like in many other developing markets, investors may not be able to diversify their portfolios optimally due to transaction costs (including liquidity-related costs such as bid-ask spreads) and information gathering and processing costs. This incomplete diversification is often combined with weak shareholder protection and corporate governance, high ownership concentration, and potential conflicts of interest between large and small shareholders. Secondly, the number of actively traded stocks in Russia has been limited, and the market weight of the oil and gas sector has been high. Thirdly, although the market infrastructure has developed (i.e., the presence of foreign institutional investors has increased together with the degree of financial integration with the world capital market), recent events, such as the Ukraine crisis and diminishing confidence in the Russian economy, have influenced the pricing of Russian stocks. Therefore, the Russian industry sectors' performance has become more sensitive to the common global market movements over time, but the allocation of financial resources has been inefficient, implying a need for stock market reforms. (Kinnunen & Martikainen 2017, 2528–2529.)

Overall, the general question of the attractiveness of the Russian equity market has provoked mixed reactions. For example, Saleem (2009, 255) states that the weak integration can offer good opportunities for international investors, whereas Castagneto-Gissey and Nivorozhkin (2016, 80, 93) emphasize that the Russian market is unpredictable, turbulent, and characterized by pervasive state control. Kinnunen (2013, 113) continues by noting that although the brief history of the Russian equity market has been characterized by high risk (measured by the substantial standard deviation of the excess market return), the market should be of potential interest for a global investor because its expected returns are driven by different factors than those affecting the investor's global portfolio.

Furthermore, relying on the optimal portfolio theory, Rockinger and Urga (2000, 457) remind that the introduction of an asset into a portfolio decreases that portfolio's risk for a given level of expected return if the asset introduced is not perfectly correlated to any of the other portfolio's assets. Correspondingly, McGowan (2011, 31–32) argues that the addition of the developing markets to the investment universe increases the efficient frontier, even when those markets are not good stand-alone investments. Finally, despite the fact that practitioners tend to promote investments in emerging markets by referring to their high returns (Kinnunen 2013, 110), foreign institutional and retail investors may be forced to tilt their Russian portfolios toward the largest and most transparent companies due to the above-mentioned indirect investment restrictions and information asymmetries (Kinnunen & Martikainen 2017, 2530). Gaddy and Ickes (2010, 291) tackle this by asserting that the aim of capital allocation is efficiency, not diversification for its own sake. Hence, diversification is a way to reduce risk, whereas specialization and comparative advantage are keys to efficiency.

With respect to previous research, an extensive literature is devoted to the issues of volatility modeling, integration, predictability, and asset pricing in emerging markets (see, e.g., Harvey 1995; Bekaert & Harvey 1997; Garcia & Ghysels 1998; Claessens, Djankov, and Klingebiel 2000) and solely in Russia (see, e.g., Hayo & Kutan 2005; Jalolov & Miyakoshi 2005; Goriaev & Zabotkin 2006; Saleem & Vaihekoski 2008; Saleem 2009 & 2011; Fedorova & Pankratov 2010; Saleem & Vaihekoski 2010; Babecký et al. 2013; Kinnunen 2013; Vizgunov et al. 2014; Kinnunen & Martikainen 2017). In addition, risk management and Value-at-Risk modeling in the Russian stock market are discussed, e.g., in Andreev et al. (2010) and Logoveev and Cherinko (2015), whereas market efficiency is considered, e.g., in Buklemishev and Maliutina (1998), Rockinger and Urga (2000), Hall and Urga (2002), Abrosimova, Dissanaike and Linowski (2005), McGowan (2011), Compton, Kunkel, and Kuhlemeyer (2013), Davydov and Vähämaa

(2013), Saleem (2014), Berezinets et al. (2017), and Caporale and Zakirova (2017). Finally, various trading rules are tested, e.g., in Chong, Cheng, and Wong (2010), Luukka et al. (2016), and Pätäri et al. (2017).

The most relevant findings of the previous studies can be briefly summarized as follows: Firstly, Hayo and Kutan (2005) analyze the impact of oil prices, news, and international financial market developments on daily returns on Russian stock and bond markets and detect that oil price movements can significantly destabilize Russian markets. According to Saleem (2011), there is also clear evidence of time-varying correlations between the Russian bond and stock market and that both asset markets exhibit positive asymmetries. Fedorova and Pankratov (2010, 167–168) continue by evaluating the influence of macroeconomic factors on the Russian stock market and show that there is a heavy dependence of the dynamics of the Moscow Interbank Currency Exchange (MICEX) index primarily on the oil price and US dollar exchange rate. Since Russia is a major oil-exporting country, an increase in the Brent oil prices has, rather unsurprisingly, a favorable effect on stock market growth and thus on national economic development. However, the dependence of the dynamics of the MICEX index on the US dollar exchange rate is inverse, i.e., the stock index value decreases, and the rate of economic growth falls on average when the US dollar exchange rate increases (see also Vizgunov et al. 2014).

Furthermore, Saleem and Vaihekoski (2008) report that even though the Russian stock market seems to be partially segmented, the global equity market influences significantly on its performance. They also find support for the pricing of the currency risk in Russia (cf. Lozinskaia & Saltykova 2019, 10). Jalolov and Miyakoshi (2005, 377), in turn, note that international investors must monitor German signals since the spillover effects from Germany have stronger impacts on Russian stock returns than those of the United States. On the other hand, Babecký et al. (2013) find evidence for the gradually increasing process of stock market return convergence between Russia and China.

Kinnunen (2013) complements the previous studies by investigating whether the relevance of a conditional multifactor asset pricing model and autocorrelation in predicting the Russian aggregate stock return fluctuates over time. He finds that the source of return predictability varies considerably with information flow, measured by volatility and trading volume. More precisely, autocorrelation increases during periods of low information flow, whereas conditional exposure to the local market risk and changes in oil price influence the expected return on the Russian stock market during periods of high information. However, the lagged global stock market factor and currency returns seem to have an insignificant influence. According to Kinnunen (2013, 107, 116), the predictability of the Russian stock market return is at a high level in general, and autocorrelation can be used to form profitable trading strategies. Finally, Kinnunen and Martikainen (2017) investigate the relation between expected returns and idiosyncratic risk in the Russian stock market and find that this unique asset-specific risk affects stock returns and is economically important in explaining cross-sectional variation in industry-level returns.

Considering market efficiency (see Fama 1965), Hall and Urga (2002) apply a test of changing market efficiency to the Russian stock market index returns from 1995 to 2000 and find that the market was initially inefficient but became efficient after around two and a half years. The tendency towards market efficiency can be explained by the fact that, during the time of the study, the Russian market was new, trading was very thin, firms' disclosure practices were very limited, and there were institutional barriers to trade (Hall & Urga 2002, 2; see also Buklemishev & Maliutina 1998). In contrast, Rockinger and Urga (2000) are not able to draw any conclusions concerning the efficiency of the Russian market, but they note that a constantly significant level of predictability characterizes the market. Hence, the market may be predictable yet efficient, and the variations of predictability can be explained by market illiquidity or by political and economic risk premium (Rockinger & Urga 2000, 470–471).

The subsequent study by McGowan (2011) evaluates the weak-form efficiency of the Russian stock market using the Russian Trading System (RTS) index for the period from 1995 to 2007. He notes that, even though the results are mixed, the RTS index is generally weak-form efficient, particularly at the end of the testing period. This can be explained by the increased international interest in the Russian stock market right before the financial crisis of 2007–2008. In line with McGowan (2011), Davydov and Vähämaa (2013, 149), who examine the relationship between stock returns and the sources of corporate debt, also find evidence that at least the weak form of stock market efficiency holds in Russia. On the other hand, Saleem (2014) utilizes the volatility modeling approach to study the temporal dependencies in the major sectors of the Russian equity market. Using the data from 2004 to 2013, he finds evidence of long memory, implying that the market sectors are weak-form inefficient. This implication of a weakly inefficient stock market is also confirmed by Abrosimova et al. (2005). However, the authors emphasize that, whilst their serial dependence results provide some limited evidence of short-term market predictability, there is not enough evidence to suggest that it would lead to a profitable trading rule once transaction costs and risk are considered.

Compton et al. (2013), in turn, seek to find evidence of calendar anomalies in the Russian stock and bond markets during the first decade of the twenty-first century. Based on the authors' findings, there is strong support for a turn-of-the-month effect in the Russian stock and bond markets, as well as strong evidence of weekday seasonality (but no Monday effect) in the Russian bond market. Compton et al. (2013, 1150) conclude that although the Russian Federation has been one of the most rapidly developing markets in Eastern Europe, the country's stock and corporate bond markets have not been as efficient as the markets in the United States. However, Caporale and Zakirova (2017) highlight the importance of considering transaction costs (proxied by the bid-ask spreads) because once these are included in the analysis, calendar anomalies disappear. According to Caporale and Zakirova (2017, 108), the absence of exploitable profit opportunities based on calendar anomalies suggests that the Russian stock market might be informationally efficient.

The opposite conclusion is made by Berezinets et al. (2017), who investigate the average reaction of the Russian stock market to dividend surprises by testing the dividend signaling theory. They find evidence of market inefficiency as there are abnormal stock returns around the dividend announcement dates. According to the authors, the presence of excess returns points out that not all publicly available information is instantly incorporated into the stock prices. Hence, it might be possible to benefit from abnormal price movements in the short term when trading prudently. (see Berezinets et al. 2017, 171, 176.)

When it comes to the profitability of various trading strategies, Chong et al. (2010) find that among the BRIC countries, trading rules based on technical indicators, namely the simple moving average, momentum, relative strength index, and moving average convergence divergence, are the most profitable in the Russian stock market. Moreover, applying these simple strategies to the Russian market remains profitable even after the relatively high transaction costs are considered (Chong et al. 2010, 238). In the same way, Luukka et al. (2016) use technical analysis to examine the profitability of Dual Moving Average Crossover (DMAC) trading strategies in the Russian stock market over the 2003–2012 period. Their results show that the best DMAC rules of the in-sample period can also outperform the passive buy-and-hold benchmark portfolio during the subsequent out-of-sample period and that the outperformance is mostly attributable to the strategies' ability to determine trends consistently (see Luukka et al. 2016, 2435, 2446; Rozhkov 2005, 65).

Interestingly, since the performance gap between the best past trading strategies and their passive benchmark portfolio does not seem to have narrowed over time in the Russian stock market, the findings of Luukka et al. (2016, 2447) contradict the assumption that the trading strategies' prediction power fades over time. Instead, their evidence is in line with the adaptive markets hypothesis of Lo (2004, 23): It states that market efficiency may not follow a systematic trend toward greater efficiency, as argued by proponents of the efficient market hypothesis, but can be context-dependent as well as dynamic. Further confirmation for such a conclusion is provided by Pätäri et al. (2017, 315–316), who continue the study of Luukka et al. (2016) and show that there is time variability in relative performance between trading portfolios of individual stocks and index trading. Finally, even though these trading strategies are simple and are not based on portfolio optimization, they provide preliminary indications of potential investment opportunities in the Russian equity market. Naturally, this should bode well for our research.

1.2 Aim and structure of the thesis

In this thesis, we seek to answer the following research question: Is it possible to invest profitably in the emerging and volatile Russian stock market while at the same time taking uncertainties, risks, and dependencies into account? We tackle this question in five parts, each of which constitutes a separate entity. In the first part, we focus on classical and robust mean-risk portfolios. We begin by examining whether the performance of the expected return maximization strategy, which is prone to estimation errors, can be improved by maximizing the mean return per unit expected shortfall and then requiring that each asset contributes equally to the portfolio's downside risk. Next, we investigate whether the use of robust covariance estimators can improve the performance of minimum variance portfolios. Thirdly, we analyze how portfolio solutions produced by robust optimization techniques differ from the Markowitzian mean-variance portfolios.

In the second part, we follow the methodology proposed by Boudt, Lu, and Peeters (2015) and study whether the out-of-sample performance of portfolio allocation schemes can be improved by estimating the higher-order comoments under a statistical factor model as opposed to using the traditional sample estimator approach. The underlying idea here is that although skewness and kurtosis can help characterize non-linear dependence in multivariate analysis, their inclusion in the portfolio optimization process may lead to a poor outcome due to the curse of dimensionality. In this context, we further compare the risk/reward profiles of two strategies based on non-Gaussian objective functions. In the first case, we minimize the risk measured by the modified expected shortfall, and, in the second case, we maximize the expected value of the fourth-order expansion of the constant relative risk aversion (CRRA) utility function.

In the third part, we utilize the ordering information framework proposed by Almgren and Chriss (2004; 2005) to mitigate the estimation risk related to the first moments. As our focus shifts towards tactical asset allocation, we analyze a strategy where the relative ranking of expected asset returns is based on the short-term momentum assumption. We also investigate whether this approach can outperform a similar asset sorting strategy based on the entropy pooling (EP) concept proposed by Meucci (2008b), as well as a quadratic utility maximization strategy using sample estimates for the first and second moments.

In the fourth part, we first study whether we can construct a profitable strategy by applying the Black-Litterman approach to the one-step-ahead forecasts derived from a vector error correction model (VECM). We then analyze how the optimal portfolio allocation changes when the latest return forecasts are employed in Meucci's (2006) copula opinion pooling (COP) framework. Secondly, within the entropy pooling framework, we examine the performance of a tangency strategy in which subjective views are formed with respect to the conditional volatilities for assets, i.e., we employ one-step-ahead predictions deduced from a generalized autoregressive conditional heteroskedasticity (GARCH) model.

Finally, in the fifth part, we examine a TAA-related wealth protection strategy proposed by Pfaff (2007; 2016). The purpose of this application is twofold: First, we demonstrate that instead of using options, a portfolio floor value that may not be violated can be defined by means of a linear program. Second, as return maximization is used as the target, we study whether the utilization of an autoregressive integrated moving average (ARIMA) model in producing return expectations can improve the portfolio's risk-adjusted returns compared with the solution of an equal-weighted strategy.

We point out that even though Russia has been previously investigated in the financial literature mainly in the context of asset pricing models and rules-based technical analysis, there have not been, to the best of our knowledge, comprehensive performance surveys of advanced portfolio optimization strategies focusing on the country's stock market. Our research intends to cover this gap. We carry out all the computations using two programming languages, namely R (see R Core Team 2018) and MATLAB® (see MATLAB 2020). In order to save space, we omit reporting the hundreds of lines of code needed to write these applications. However, all the algorithms are available for inspection from the author upon request. Throughout our empirical study, we take the perspective of a Russia-based investor by using ruble-denominated returns and do not consider currency hedging (see, e.g., Celebuski, Hill & Kilgannon 1990; Moosa 2003; Kinlaw & Kritzman 2009; Schmittmann 2010; Boudoukh et al. 2019; Arruda, Bergeron & Kritzman 2021 for more discussion on optimal currency exposures in international equity portfolios).

In the next subchapter, we briefly describe the development of the Russian stock market. The rest of the thesis is organized as follows: Chapter 2 is dedicated to the theory of portfolio selection and portfolio risk management. More precisely, in Subchapter 2.1, we present the main features of the Modern Portfolio Theory (MPT), whereas, in Subchapter 2.2, we review the issues highlighted in the literature related to the classical mean-variance optimization. The Black-Litterman model with its extensions and modifications as a special remedy to these optimization problems is discussed in Subchapter 2.3. In Subchapter 2.4, we focus on the momentum anomaly often found in financial markets. Finally, we end this chapter by studying the properties of tail risk measures widely promoted in the literature, namely Value-at-Risk (VaR) and Expected Shortfall (ES). We also briefly discuss how these risk measures can be combined with volatility and dependence modeling to create dynamic risk assessment systems. The data and descriptive statistics are reported in Chapter 3.1, and the rest of Chapter 3 explains the methodology behind our portfolio strategies divided into five subchapters. Correspondingly, Chapter 4 presents the empirical findings and conclusions for each application in five subchapters. Finally, in Chapter 5, we summarize our results.

1.3 Development of the Russian stock market

Only a few studies exist that consider the historical development of the Russian stock exchange. For example, the papers by Lucey and Voronkova (2005) and Anatolyev and Shakin (2007) are outdated due to the rapid changes that have happened in the Russian economy after the 2007 global financial crisis. This chapter strives to bridge the gap in the literature by providing a brief overview of the main events in the Russian capital market, starting from the collapse of the Soviet Union and ending to the aftermath of the Ukrainian crisis. At the end of this chapter, we briefly speculate on the future of the Russian financial system.

To begin with, the development of the Russian stock market has mirrored the country's integration into global financial markets, as well as swings in investor perceptions. Except for Estonia, the Russian Federation has probably been the most successful of the 15 former Soviet Union republics in improving its economy and developing functioning financial markets (Compton et al. 2013, 1139). During a severe economic crisis in late 1994, the country had to create the stock market as a part of its transition from the planned system to the market economy. At the end of 2005, Russia's stock market had become one of the largest emerging markets in the world. This growth was driven both by the expansion in the number of stocks and by high returns. However, the market has also suffered from many structural deficiencies, such as illiquidity, high volatility and concentration, and low diversification potential. (Goriaev & Zabotkin 2006, 380–381.)

Multiple stock exchanges had been founded in Moscow already at the beginning of the 1990s, even before the Russian Government had issued any legislative acts (see Hall & Urga 2002, 7). Nevertheless, the evolution of the Russian financial system was mainly due to the increasing foreign exchange trading on the Moscow Interbank Currency Exchange (MICEX). The significant expansion of the scale of US dollar operations was caused by the liberalization of foreign economic relations and the rejection of a plurality of exchange rates. On the other hand, the beginning of the privatization process, which was accompanied by the 'emission' of 150 million state-securities vouchers and by voucher auctions in 1993–1994, created conditions for the emergence of one more segment in the system of financial markets, i.e., the market for stocks of privatized companies. (Iakovlev & Danilov 1997, 47–49.)

As the initial phase of Russia's mass privatization was ending in late 1994, an appreciable and quite prolonged slowdown of inflation triggered an influx of foreign portfolio investments estimated at four to six billion dollars (Iakovlev & Danilov 1997, 49). In order to consolidate regional securities markets into an organized security industry, the electronic Russian Trading System (RTS) was launched in mid-1995 (Hall & Urga 2002, 8; Anatolyev 2008, 57). RTS was also the initial platform for the Russian derivatives market. MICEX, in turn, expanded into equity trading in mid-1997, offering to trade in government bonds, foreign currency, derivatives, as well as corporate bonds and stocks. Both the MICEX and RTS calculated the capitalization-weighted composite indices of the 30–50 most liquid stocks of Russia's largest companies representing approximately 80 percent of the country's stock market. However, the former index has been denominated in rubles, whereas the latter has been US dollar-denominated. Nearly 15 years later, in December 2011, these two largest and centralized trading platforms merged, and the formed exchange was named the Moscow Exchange MICEX-RTS (nowadays simply MOEX). (Compton et al. 2013, 1143.)

During 1995, several attempts were also made to set up a market regulatory system. Many decrees appeared aiming to create the institutions needed to operate the market and protect shareholders' rights. In April 1996, the signing of the Law on the Securities Market established the role of the Federal Commission on Securities Market (FCSM) of Russia as the principal regulator. (Hall & Urga 2002, 7.) However, various scams and low confidence in financial companies that accepted deposits from the population and were active in the bearer stocks market led to the transfer of funds to the currency market. The unstable exchange rate was used for speculation by foreign portfolio investors, most of whom exited the Russian market after the First Chechen War began. Furthermore, the excessive participation of the state in trade did not increase stability but only generated new types of risks due to the vagueness of the decisions of the Central Bank of Russia (CBR) and the Ministry of Finance. Nevertheless, one could have achieved profitability of 15–20 percent per month on investment, even though this was possible only for a person who was present on the trading floor and had not paid a broker commission on every transaction. Otherwise, the profits would have gone to paying transaction costs. (Iakovlev & Danilov 1997, 51, 61–62; see also Ratinov 1997.)

Although the RTS index displayed an impressive 94 percent growth from the beginning of 1997 until October, positive tendencies in the stock market were taking place against the background of poor fundamentals in the Russian economy. Several critical issues (such as the banking system vulnerability, budget crisis, and a high value of short-term government liabilities relative to the central bank reserves) were exacerbated by the instability of the international financial markets, mainly by events in Southeast Asian markets in Fall 1997. Under these circumstances, foreign investors began to sell government and corporate bonds. Moreover, increased demand for foreign currency caused a deep decline in central bank reserves. Eventually, these events were reflected in the falling stock market: By January 1998, the RTS index had plummeted by 50 percent. (Lucey & Voronkova 2008, 1306.)

The decisive intervention from the Central Bank in January 1998 initially prevented a currency crisis from escalating and helped the equity market to recover. However, already in March 1998, political instability and the fall of the oil price, with the impact on the State budget, negatively affected the equity market. Pressure from the exchange rate forced the Central Bank to increase interest rates to more than 100 percent at the beginning of the second quarter of the year. (Hall & Urga 2002, 8, 11–12.) Between March and May in 1998, there was a further 20 percent decline in stock market prices. Despite financial aid provided by the International Monetary Fund (IMF) and International Bank for Reconstruction and Development (IBRD) in July, a further decline in security prices occurred as Russian banks and foreign investors continued their selling spree. (Lucey & Voronkova 2008, 1306–1307.)

In August 1998, the Russian Central Bank's international exchange reserves were exhausted, and the federal budget could not service the government's debt. This caused the Russian government to declare a default on its bills and bonds, including a default on ruble-denominated debt instruments that at some point were perceived to be the 'risk-free asset' in the economy (Santos 2003, 166). In addition, the Russian Central Bank was forced to stop adhering to a fixed exchange rate regime. As the country's financial system became paralyzed, there was also a moratorium on debt payments, devaluation of the ruble of more than 50 percent, and a halt in foreign exchange trading on MICEX. Several of the largest Russian commercial banks were eliminated, and many individuals and non-financial companies lost their savings. Between August and September, the RTS index collapsed by almost 70 percent. By the end of the year, the stock market was virtually nonexistent, despite an official figure of 237 listed companies but with a capitalization of only 20.6 billion US dollars. In addition to foreign capital outflows, oil prices had declined to 10–11 US dollars per barrel. (Hall & Urga 2002, 8, 11–12; Smirnov 2015, 142; see also Chiodo & Owyang 2002; McAllister & White 2011, 477–479.)

Stock returns recovered strongly from the huge losses already in the following year, even though international interest in the Russian market was declining, and trading activity, which had fallen by 84 percent since 1997, was in record-low levels. According to Lucey and Voronkova (2008, 1307), low turnover created pre-conditions for speculative growth of the market that ascended to 194 percent in 1999 and made the RTS the fastest growing market in the world. Several other factors may also explain this unexpected rise. These include the debt restructuring, the dramatic rise of oil prices, the expanding trade surplus and positive industrial production, the agreement between the Government and the IMF to repay old loans with new ones, as well as the restructuring of the GKOs (i.e., short-term zero-coupon government bonds or Treasury-bills). (Hall & Urga 2002, 12–13.) In 2000, despite the rapid growth of the Russian economy, the stock market showed a disappointing performance as RTS declined by 20 percent. This was mainly caused by a decline in prices of Russian blue chips (mostly oil companies) that depended heavily on the dynamics of the oil prices. On the other hand, the improving political and macroeconomic situation helped revive investors' interest and boost turnover, which more than doubled in 2000. (Lucey & Voronkova 2008, 1307.)

After Vladimir Putin was elected president in March 2000, the perception of binary political risk was no longer driving the market as decisively as before. However, the equity market's history under the control of President Putin has not been trouble-free: In 2003, the political risks of investing in the Russian market increased once again. This was due to the conflict between the oil company Yukos and the government, which eventually led to the imprisonment of the company's chief executive, Mikhail Khodorkovsky. The market reacted to this event with a 25 percent decline during October 2003. Arguably, investors interpreted it as a signal about the toughening of the government policy towards the business community. (Goriaev & Zabotkin 2006, 389–390; Lucey & Voronkova 2008, 1307, 1320; see also Goriaev & Sonin 2005.)

While Russia's real GDP grew at about five percent during 1999–2005, the stock market grew at a compound annual rate of over 50 percent (Black et al. 2006, 364). At the end of 2005, the RTS index achieved the level of 1125.6, more than an 11-fold increase from its starting value. However, the degree of volatility had also been high: Firstly, the annualized standard deviation of the RTSI's weekly returns for the period from its start till the end of 2005 was 7.3 percent. Secondly, the cross-sectional variance of stock returns in Russia was also considerable, meaning that stock picking and the understanding of firm- and industry-specific risk were not of secondary importance. Thirdly, the relationship between the RTS index and global equity indices varied greatly over time, implying a smaller dominance of the country risk than in the early years of the market. (Goriaev & Zabotkin 2006, 386–387; see also Anatolyev 2008, 60–61, 65.)

The number of listed companies was over 350 on the RTS exchange in 2001, but most stocks were traded irregularly. Since 2000, the number of stocks with trades registered on any given day varied between 20 and 60. This number increased by around 50 percent to almost 90 in late 2005. The rapid expansion of the domestic market was partly due to new companies going public. The Russian stock market saw 13 initial public offerings and six secondary offerings between 2004 and 2005. The offerings coming from the retail, food, media, metal, and mining sectors further improved the diversification opportunities for investors operating in Russia. Another trend was the increasing accessibility of the Russian stock market to international investors and large regulated institutional investors. (Goriaev & Zabotkin 2006, 382–385.) In 2006, the Russian government and the Bank of Russia lifted nearly all the restrictions on capital transactions, which lowered visible risks for foreign investors (Kudrin & Gurvich 2015, 34).

Before the global financial crisis of 2007–2008, the Russian stock market was at an all-time high. The significant inflow of foreign capital was caused by the strengthening ruble and the growth in production and oil and gas revenues (Kudrin & Gurvich 2015, 36–37). From the equity market perspective, companies in the commodity-exporting sectors, which dominated the economy, created value, and new businesses – especially in service and consumer-oriented sectors – emerged and entered the public capital markets.

Moreover, the consumer goods sector, which was driven by quickly growing consumer demand in Russia, appeared to be the least sensitive to global trends. (Goriaev & Zabotkin 2006, 390, 394, 396; see also Gaddy & Ickes 2010, 284–287, 290.)

During 2007, construction, electricity, consumer goods, and banking industries raised nearly 30 billion US dollars in over 30 public offerings. Russia's market capitalization at the time was comparable to that of China and India, with only 190 companies trading on Russian exchanges. Furthermore, although the overall turnover of shares remained low, the involvement of foreign investors on the MICEX exchange increased from 15 to 25 percent of the daily trading volume. (Compton et al. 2013, 1139-1140.) At the same time, a system was established for accumulating state reserves, borrowing portfolio capital abroad, and exporting private capital. Hence, in the absence of a national investment process and mechanisms for reinvesting export earnings, evolving challenges with portfolio borrowing abroad destabilized the savings system. (Grigoriev & Salikhov 2009, 47–48.)

Starting from August–September 2007, Russian banks faced a liquidity problem and a foreign credit crunch. Simultaneously, at the peak of the international boom, domestic companies continued to be active in mergers and acquisitions, putting up blocks of shares as collateral to finance these deals. The market indices rose until May 2008, but such a short-sighted approach to finance large-scale transactions became evident in the fall, as margin calls put several companies on the verge of a change in ownership. Eventually, the government decided to redeem major owners from Western creditors. (see Grigoriev & Salikhov 2009, 48–51; Andriushin & Burlachkov 2009, 58–59.) Moreover, the money saved during the time of high oil prices was swiftly spent on supporting banks' liquidity and protecting the ruble exchange rate (see Tabata 2009, 685–686; Kudrin 2009, 21).

The dependence on foreign short-term lending made the domestic economy vulnerable to external shocks. As an anti-crisis measure, the authorities injected large sums into the economy. However, the magnitude of the anti-crisis program (as much as 1.5 percent of GDP) and the quick drawdown of currency reserves led to unrest in international financial markets and weakened the positions of Russia and Russian issuers on international trading floors. (Grigoriev & Salikhov 2009, 51.) Further, the Central Bank of Russia owned about 100 billion US dollars of mortgage-backed securities during the crisis (Davydov & Vähämaa 2013, 149). Overall, the Russian stock market had one of the largest declines, even among developing countries. For example, the MICEX price index plummeted by about 74 percent from 1956.14 (recorded on May 19, 2008) to 513.62 (on October 24, 2008), in as little as five months (Iwasaki 2014, 178).

The global financial crisis caused severe damage to the Russian industry, and an abnormally high number of firms went bankrupt. For instance, the Russian manufacturing industry, which is sensitive to global economic conditions, showed signs of a change earlier in 2008. Then, in 2009, its value-added production quickly shrank by 14.9 percent in real terms compared to the previous year. (Iwasaki 2014, 179.) As a result, the manufacturing index fell by over 90 percent from peak to trough after the market's mid-2008 peak (Gaddy & Ickes 2010, 290). On the other hand, the impact of the crisis was not uniform at the sectoral level, suggesting potential for diversification of risk across sectors (Babecký et al. 2013, 16). The overall market capitalization of Russian companies fell 74 percent, from 1.5 trillion US dollars in 2007 to 397 million dollars in 2008. However, by the end of 2010, the market had recovered 40 percent of its value and ended the year at over 1 trillion dollars. (Compton et al. 2013, 1140.) Furthermore, although the share of state-controlled firms in the total capitalization of Russian companies had increased from 24 percent to 40 percent between 2004 and 2007, the long-term reduction in the stock market's total capitalization was accompanied by a reduction in the share of the state-owned enterprises (Abramov, Radygin & Chernova 2017, 9–10; see also Nesvetailova 2016).

With respect to dividend policies, Ankudinov and Lebedev (2016) find that during the global financial crisis, the dividend payments of state-controlled firms decreased more significantly than those of privately-owned companies. Historically, dividends per share have also been volatile, but the shareholders typically perceive stocks of the companies from the oil and gas and from the utility sectors to be 'cash cows,' from which large dividend yields are expected. On the other hand, the dividend payout ratios of Russian companies have generally been substantially lower than in both developed and developing markets. (see also Berezinets et al. 2017, 170–171.) Liljeblom and Maury (2016) show further that in Russia, a significant increase in dividend payout levels coincides with improvements in legal shareholder protection.

After the crisis, there was a positive growth trend in the average dividend yield, and by 2014, the average annual dividend yield reached the level of 6.1 percent. This was, in some respects, a consequence of the new regulation of dividend payouts (25 percent of net earnings) for state-owned companies, which came into force in November 2012. In 2015, following the results of 2014, the total dividends announced by Russian companies increased in absolute terms by over 20 percent, up to roughly 14.6 billion US dollars despite the recessionary trends. However, as a sign of high concentration, ten companies accounted for 88.1 percent of the total dividend payments in 2013. (Berezinets et al. 2017, 161–163; see also Koesterich 2014.)

The crisis of 2008 also caused changes to the rules of short selling on the Russian stock markets. The country had started to regulate short-selling explicitly since 2002, which stemmed from the default and collapse of the Long-Term Capital Management (LTCM) hedge fund in 1998. In September 2008, a Russian governmental body, the Federal Financial Markets Services (FFMS), announced that it would ban margin trading and short selling altogether to prevent further speculations on sharply decreasing stock prices. This position was further affirmed by the closure of the two leading exchanges for two days.

(Pekarek & Meseha 2012, 372.) During September 2008 and June 2009, as short sales were prohibited, the volatility of the MICEX index decreased considerably. Moreover, the Sharpe ratio and kurtosis were low, while skewness was high. Short-selling was allowed temporarily in June 2009 to improve market liquidity since many market operators and institutional investors had abandoned the Russian market. The remaining speculators had relied on complex arrangements to overcome the short-selling ban, and some brokers were even accused of trading the clients' securities without authorization. The short-selling constraint was eventually lifted at the end of 2010. (Kudrov et al. 2012, 392–394.)

Currently, the Russian authorities require that investors provide their brokers with collateral in cash, cash equivalents, or other securities. Furthermore, brokers are entitled to short sell only quoted securities in limited quantities. The exchange forms quarterly a list of the most liquid securities that are allowed for short selling in the next quarter according to the liquidity rating rules set by the FFMS. Securities provided as collateral need to fulfill the exact liquidity requirements, and limitations are set for short sales if the margin falls below the allowed minimum of 50 percent. Also, a broker is not allowed to execute short sale orders if the price is less than the closing price for the previous trading day minus three percent. Average costs for short-selling have been high, and some brokers in Russia do not even offer short-selling due to its high risk. (Kudrov et al. 2012, 388–393; see also Compton et al. 2013, 1142 for more details on institutional changes.)

As a sign of the recovering Russian market, the RTS and MICEX indices peaked in April 2011, although they did not reach their pre-crisis maximum values. After that, the market demonstrated a decreasing trend, caused by the stagnation of the Russian economy (Voskoboynikov 2017, 357; see also Delcoure 2007; Compton et al. 2013, 1141). Furthermore, between 2012 and 2013, high inflation (6–8 percent per year) prevented the Central Bank from lowering target interest rates (Smirnov 2015, 143), and the attempt to stabilize the deteriorating ruble exchange rate forced it to reduce the international reserves (Ono 2017, 323). Liquidity also remained low, which can be considered as one of the reasons for the undervaluation of many Russian companies' shares. The downfall deepened against the background of the events of 2014 when the collapse of the Russian stock prices occurred. During 2014, oil prices plunged, the ruble crashed, and the total market capitalization of Russian companies dropped by roughly 50 percent. As a result, the performance of the Russian stock market was the worst in the world in 2014. (see Berezinets et al. 2017, 159, 161; Veebel & Markus 2016, 135, 138–139.)

The beginning of the Ukrainian crisis, which can be traced back to March 2014, represents a new phase in the Russian economy. The United States, the European Union, as well as other countries and international organizations imposed several rounds of ever-tighter sanctions against individuals and businesses from Russia. For example, Russian state-controlled banks (including *Bank of Moscow*, *Gazprom Bank*, *Sberbank*, and *VTB*

Bank) were excluded from raising long-term loans, the access to international debt markets was restricted for several major publicly owned companies (including oil and gas companies *Rosneft*, *Transneft*, and *Novatek*), and an EU-US ban targeted exports of some oil industry technology and services. Moreover, citizens or residents of America were banned from dealing with the Russian banks. EU nationals and companies, in turn, were no longer able to buy or sell *new* bonds, equity, or similar financial instruments issued by five major state-owned Russian banks, three major Russian energy companies, and three major Russian defense companies (see Council Regulation (EU) No 960/2014; Council Regulation (EU) No 1290/2014). (Nivorozhkin & Castagneto-Gissey 2016, 26–27; see also Gurvich & Prilepskiy 2015, 360–362; Bagheri & Akbarpour 2016, 93; Veebel & Markus 2016, 132–134, 137.)

By focusing on the US dollar-denominated MSCI Russia index, Nivorozhkin and Castagneto-Gissey (2016, 26–27) argue that since the negative monthly returns in March and in April were followed by positive returns in May and in June before turning negative for the rest of the year, the effect of the sanctions on the Russian stock market was limited. Alternatively, the investors might have been anticipating the consequences of the crisis and thus the effect of sanctions already prior to March 2014. In any case, the stock market fall was dramatic after November 2014, returning negative 80 percent in mid-December from the start of the year and ending the year 62 percent down. In 2015, the recovery of the Russian stock market, which lasted until mid-May, made it for a moment one of the best-performing equity markets in the world. However, by the end of summer 2015, the gains effectively disappeared as the market remained about 52 percent down relative to the start of 2014.

Simultaneously, there was a significant decrease in the interdependence between the Russian equity market returns and returns in most major developed, frontier, and emerging markets. In contrast, the strongly varying and often negative correlation across Russian sectoral stock indices in 2013 changed to uniformly positive co-movement across virtually all sectors in the crisis period. (see Nivorozhkin & Castagneto-Gissey 2016, 23–24, 31–35, 39.) Furthermore, a statistically significant increase in volatility for most sectoral indices in the aftermath of the crisis is found by Ankudinov, Ibragimov, and Lebedev (2017). In contrast, Hoffmann and Neuenkirch (2017, 66–68) observe a significant leverage effect in the stock market as negative innovations led to higher volatility than did positive ones. Finally, regarding shock spillover characteristics, Schmidbauer et al. (2016, 497, 506) argue that large-scale sanctions increased Russia's importance as a propagator of volatility shocks in the network of global stock markets. This finding of volatility transmissions is also in line with the study by Castagneto-Gissey and Nivorozhkin (2016).

Despite the adverse circumstances, the ruble devaluation and the falling fertilizer and fuel prices created an attractive pricing environment for the Russian agriculture sector. In 2016, Russia became the world's largest exporter of grains, and while agriculture remains far below oil and gas, this sector has bypassed arms sales to become Russia's second-biggest exporter. The country's agriculture boom shows that, despite sanctions and the poor state of East-West relations, there are niches of value and opportunity to be found in the Russian market. For example, after March 2014, the fertilizer company *PhosAgro* benefitted from the 16 percent increase in the consumption of crop chemicals (against the global growth of 2.2 percent), leading to an 85 percent rise in the ruble-dominated share price in 2016. Similarly, the shares of *Acron*, another mineral fertilizer producer, had tripled in that period. Finally, due to the increased domestic production, shares in *Cherkizovo*, the meat processor, jumped 63 percent, well above the 45 percent increase in the MICEX index. (Buckley, 2017.)

During 2016, both the basic material sector and wholesale trade began to indicate positive trends, and, by the beginning of 2017, the economic recession had effectively ended. However, the end of the depression and anti-crisis policy had not equated to resolving structural problems in the Russian economy. (Mau 2017, 117–118, 123.) Indeed, on a larger scale, the future of the Russian equity market appears gloomy, as any positive driver for economic growth is not obvious (Smirnov 2015, 144). It is also difficult to forecast when the sanctions will be lifted, and nothing promises that the country's economy could develop without a constant inflow of oil and gas rents. Although the long-term trend for oil is generally described as a slowly growing trend, massive cyclic fluctuations are superimposed on it. (Kudrin & Gurvich 2015, 40, 42; see also Simola & Solanko 2017; MacIntyre & Hess 2019.) The risk of long-term stagnation is also increased by high inflation (see, e.g., Mau & Ulyukaev 2015; Akindinova, Kuzminov & Yasin 2016), the pressure to raise taxes (Kudrin & Gurvich 2015, 50–51), inefficiencies in the country's financial sector (Karas, Schoors & Weill 2008; Pak & Kretzschmar 2016; Ono 2017), the increasing government control (Vanteeva 2016), weak market institutions (Medvedev 2016; Berezinskaya 2017), as well as inadequate protection of ownership rights (Iwasaki 2014), and poor corporate governance practices (Black, Love & Rachinsky 2006). The stock market is also largely perceived as a venue for 'respectable gambling,' but not as a source of long-term funding badly needed in order to restart investing activities in the Russian economy (Ankudinov et al. 2017, 160).

2 THEORY OF PORTFOLIO SELECTION AND PORTFOLIO RISK

In this chapter, we discuss the modern portfolio theory and its limitations. We also consider possible ways to improve portfolio optimization and present several variants of the so-called Black–Litterman model. Fourthly, we focus on a well-documented anomaly in financial markets known as momentum. Finally, we end this chapter by illustrating different ways of modeling portfolio risk.

2.1 Modern portfolio theory in a nutshell

Modern Portfolio Theory (MPT) is a compilation of tools and techniques by which a risk-averse investor can construct an optimal portfolio. It comprises the Capital Asset Pricing Model (CAPM), the efficient market hypothesis, and various quantitative portfolio construction and optimization forms. Initially, in his seminal paper 'Portfolio Selection,' Markowitz (1952) introduced the foundations of MPT, mean-variance optimization, and mean-variance analysis. The mean-variance analysis provides a framework to construct and select portfolios based on the investments' expected performance, as well as on the investor's risk aversion. Markowitz quantified the concept of diversification by using the statistical notion of covariance between individual assets and the overall standard deviation of a portfolio. Therefore, investing the capital in assets whose returns are highly correlated is not a preferable investment strategy because if any single asset's performance is poor, it is probable that these other assets will also perform poorly. This, in turn, leads to the inferior performance of the whole portfolio. (Fabozzi, Focardi & Kolm 2010, 313–314; Carl & Peterson 2014, 16–20.)

The mean-variance analysis can also be utilized to depict the efficient frontier. When there are only risky assets, the mean-variance efficient frontier has a parabolic shape. However, when a risk-free asset is added, the efficient frontier becomes linear, forming the Capital Market Line (CML). The Markowitzian premise is that of a rational investor, who at time t, determines what portfolio of investable assets to hold for a time horizon of Δt. The investor decides on the gains and losses she will make at time $t + \Delta t$, without contemplating eventual gains and losses during or after the period Δt. At time $t + \Delta t$, the investor will re-examine the situation and decide anew. This one-period framework is commonly known as myopic (or short-sighted) behavior. (Fabozzi et al. 2010, 314.)

It is further reasoned that investors should accept a trade-off between risk and expected return (as indicated by indifference curves). The expected return of an asset is defined as the expected price change plus any additional income over the time horizon considered (e.g., dividend payments) divided by the beginning price of the security. Risk, in turn, ought to be measured by the variance of returns (i.e., the average squared deviation around the expected return). Markowitz's mean-variance approach is consistent with the expected utility maximization under certain assumptions, and it does not assume the joint normality of security returns. For any given level of expected return, a rational investor should choose the portfolio with minimum variance from amongst the set of all possible portfolios (aka the feasible set). The set of all mean-variance efficient portfolios for different desired levels of expected return forms the efficient frontier. (Fabozzi et al. 2010, 314–315; see also Markowitz 1952; 1959; 1987; 1994 for more detailed expositions.)

Figure 2.1 provides a graphical illustration of the efficient frontier of risky assets (see Würtz et al. 2015b, 22). Here, nine sectoral price indices are used to represent the Russian equity market. The daily data has been downloaded from the Moscow Exchange, and it covers the period from December 28, 2007, to September 7, 2017. These indices are *Chemicals* (MCXchm), *Consumer Goods and Services* (MCXcgs), *Electric Utilities* (MCXpwr), *Financials* (MCXfnl), *Manufacturing* (MCXmnf), *Metals & Mining* (MCXm.m), *Oil & Gas* (MCXo.g), *Telecoms* (MCXtlc), and *Transport* (MCXtrn). (see Moscow Exchange Equity Indices 2017; Subchapter 3.1.) By inspecting Figure 2.1, we can see, for example, that as a separate investment, the chemicals sector has the highest mean return ($\bar{r}$), whereas the manufacturing sector has the lowest realized return. Moreover, the oil and gas sector has the highest risk, as measured by the standard deviation ($\bar{\sigma}$) of the return. In contrast, the consumer goods and services sector has the lowest standard deviation among these risky assets. *EWP* denotes the Equal Weights Portfolio, which has been constructed such that each asset in it has the same weight.

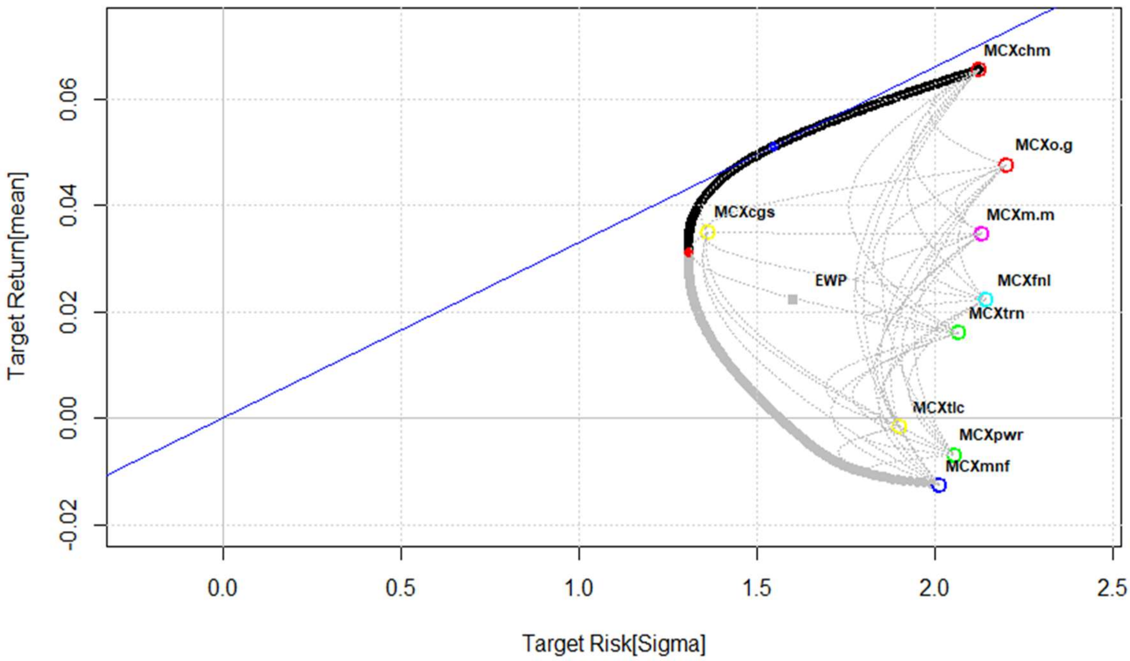

Figure 2.1 Markowitz efficient frontier and the capital market line

From Figure 2.1, it can also be observed that the locus of this set in the $\{\bar{\sigma}, \bar{r}\}$-space are hyperbolas. The set inside the hyperbola is the feasible set of mean/standard deviation portfolios, and the two borders are the efficient frontier (the black upper border) as well as the minimum variance locus (the grey lower border). The solid red dot, which divides this curve into two parts, has the smallest risk on the efficient frontier, and it is called the Global Minimum Variance Portfolio (GMVP). To the right, the feasible set is determined by the envelope of all pairwise asset frontiers. Moreover, the region outside the feasible set is unachievable by holding risky assets alone, and points below the frontier are suboptimal. Thus, a rational investor will hold a portfolio only on the frontier. The straight blue line, which represents reward/risk profiles of different combinations of a risky portfolio with a riskless asset (with a certain rate of return r_f), is the capital market line. Note that here we assume a zero risk-free rate. The solid blue dot where the CML touches the efficient frontier corresponds to the optimal risky portfolio, known as the tangency portfolio point (or market portfolio, see Fama 1970). (Würtz et al. 2009, 225, 247.)

The presence of a riskless asset in a portfolio implies lending and borrowing cash at a riskless rate. At least in the US markets, one might consider the riskless asset to be a zero-coupon bond whose maturity matches the investment horizon (e.g., a short-term government bill). Alternatively, savings accounts could represent investing in an asset with a certain outcome. Thus, if borrowing (i.e., a short position) of the riskless asset is not allowed, only the line segment between the riskless asset and points in the original feasible set can be adjoined, but the line cannot be extended further. (Lyuu 2004, 462; Elton et al. 2014, 81–84; see also Damodaran 2002; 2008; 2010 for more details on estimating country-specific risk-free rates.)

The mathematics behind the mean-variance theory of portfolio selection can be briefly presented as follows. First, let us assume that there are n assets with random rates of return, $r_1, r_2, \dots, r_n$. Thus, the expected values of these returns are $\bar{r}_i \equiv E[r_i]$. If a portfolio is formed of these n assets by using (capitalization) weights $\omega_1, \omega_2, \dots, \omega_n$, the portfolio's rate of return is

$$\boldsymbol{r} = \omega_1 r_1 + \omega_2 r_2 + \cdots + \omega_n r_n \tag{2.1.1}$$

with mean $\bar{\boldsymbol{r}} = \sum_{i=1}^{n} \omega_i \bar{r}_i$ and variance

$$\boldsymbol{\sigma}^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} \omega_i \omega_j \sigma_{ij} = \sum_{i \neq j} \omega_i \omega_j \sigma_{ij} + \sum_{i=1}^{n} \omega_i^2 \sigma_i^2 , \tag{2.1.2}$$

where σ_i^2 represents the variance of r_i and σ_{ij} represents the covariance between r_i and r_j. It is clear that $\sigma_{ii} = \sigma_i^2$. Note that if the objective were, for example, to construct a minimum variance portfolio in the absence of a risk-free asset, one would then seek to minimize Equation (2.1.2). (Lyuu 2004, 458; 460–461, 463.)

The portfolio's total risk (as measured by its variance) consists of two parts. Firstly, $\sum_{i \neq j} \omega_i \omega_j \sigma_{ij}$, the systematic risk associated with correlations between the returns on the assets in the portfolio. Secondly, $\sum_{i=1}^{n} \omega_i^2 \sigma_i^2$, the specific or unsystematic risk associated with the individual variances alone. Every possible weighting scheme $\omega_1, \omega_2, \ldots, \omega_n$ with $\sum_{i=1}^{n} \omega_i = 1$ corresponds to a portfolio, with negative weights standing for short sales. Naturally, the constraints $\omega_i \geq 0$ can be added to exclude short sales. A portfolio $\boldsymbol{\omega} \equiv [\omega_1, \omega_2, \ldots, \omega_n]^{\mathrm{T}}$ that satisfies all the specified constraints is termed a feasible portfolio. If the returns of the assets are uncorrelated, i.e., $\sigma_{ij} = 0$ for $i \neq j$, the variance of the portfolio's return decreases toward zero as n increases, provided that the portfolio is well diversified. Correspondingly, it can be shown that when asset returns are correlated, and the portfolio gets larger and is well diversified, the specific risk tends to zero, whereas the systematic risk converges to the average of all the covariances for all pairs of assets in the portfolio. (Lyuu 2004, 458–459.) Markowitz (1999, 8) calls this phenomenon the law of the average covariance.

When there is a riskless asset, the minimum variance portfolios, which are now a combination of the tangency portfolio and the risk-free asset, are superior to the portfolio on the Markowitz efficient frontier for the same level of risk. Identifying the tangency portfolio is computationally straightforward as it is the feasible point maximizing $\theta \equiv (\bar{\boldsymbol{r}} - r_f)/\sigma$ (i.e., the slope of the CML). More precisely, let us assign weights $\omega_1, \omega_2, \ldots, \omega_n$ to the n risky assets such that $\sum_{i=1}^{n} \omega_i = 1$. Since the weight on the riskless asset in the tangent fund is zero, we can see that $\bar{\boldsymbol{r}} - r_f = \sum_{i=1}^{n} \omega_i (\bar{r}_i - r_f)$, and

$$\theta = \frac{\sum_{i=1}^{n} \omega_i (\bar{r}_i - r_f)}{\sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} \omega_i \omega_j}}. \tag{2.1.3}$$

It is also easy to verify that the tangency portfolio can be calculated directly from the maximal Sharpe ratio optimization problem (see Sharpe 1994). (Lyuu 2004, 463; Fabozzi et al. 2010, 323–324, 358; see also Dixit 1990 or Griva, Nash & Sofer 2009 for more discussion on solving Markowitzian and other quadratic programming problems.)

Based on the above mean-variance analysis, it is straightforward to derive the Capital Asset Pricing Model, which was initially developed by Sharpe (1964) but also independently arrived at by Lintner (1965) and Mossin (1966). By definition, CAPM is a market equilibrium model or a general equilibrium theory of the relation of prices to risk, but it is usually applied to partial equilibrium portfolios. This can create severe problems in valuation. (see Carl & Peterson 2014, 16.) Furthermore, although the CAPM has been widely used by practitioners (Leland 1999), it should be emphasized that many of its assumptions (e.g., normally distributed asset returns or quadratic utility functions) are unrealistic. It is also implied that investors view upside and downside risks with equal aversion. The list of critical remarks regarding the CAPM and its modifications is long (see, e.g., Fama & French 1992; Campbell 1996; Malkiel & Xu 1997; Miller 1999; Sharpe, Alexander & Bailey 1999; Haugen 2001; Brown & Walter 2013; Dempsey 2013; Malkiel 2019). On the other hand, despite the controversies, the CAPM framework may justify passive or index investing by assuming that assets will rise or fall in price until they are on the efficient frontier of the market portfolio (Carl & Peterson 2014, 20).

Indeed, prices should, in theory, adjust so that portfolios fall on the capital market line. Individual risky assets and inefficient portfolios, in contrast, will plot below the line. As the CML shows the return offered to compensate for a perceived level of risk, each point on the CML is a balanced market condition (or equilibrium). The slope of the CML, also known as the market price of risk, tells by how much the expected rate of return of an efficient portfolio must increase if the standard deviation of that rate increases by one unit. It can be further shown that the expected return on a portfolio is equal to the risk-free rate plus a risk premium. The risk premium, in turn, is equal to the market price of risk (as measured by the reward per unit of market risk) times the quantity of risk for the portfolio (as measured by the standard deviation of the portfolio). (Fabozzi et al. 2010, 326; see also Fama 1968; Black 1993; Gray 1993 for more discussion on estimating expected return and risk premium.)

Although the capital market line relates the expected rate of return of an efficient portfolio to its standard deviation, it does not show how an individual asset's expected rate of return relates to its risk. Therefore, we need the following theorem: If the market portfolio M is efficient, the expected return $\bar{r}_j$ of any asset j satisfies $\bar{r}_j - r_f = \beta_j(\bar{r}_M - r_f)$, where $\beta_j \equiv \sigma_{j,M}/\sigma_M^2$ and $\sigma_{j,M} \equiv \mathrm{Cov}[r_j, r_M]$. In other words, beta (not volatility) measures an asset's risk, and it can be estimated by regressing the excess return on the asset against the excess return on the market. The CAPM formula represents a linear relation between beta and the expected rate of return for all assets, and when this relationship is plotted, it forms the Security Market Line (SML). The market is the point at $\beta = 1$, and the method of beating the market is to assume greater risk (i.e., beta). Note that the same arguments apply even if there is no riskless asset. (Lyuu 2004, 465–466.)

Finally, to save space, we refer the interested reader to the following studies regarding investment management. Firstly, risk-adjusted performance evaluation is discussed in Modigliani and Modigliani (1997), whereas Gollinger and Morgan (1993) and Gyourko and Keim (1993) demonstrate that the framework of modern portfolio theory can also be applied to commercial bank loans and real estate, respectively. Jin, Markowitz, and Yu Zhou (2006) show that the mean-semivariance efficient strategies, which correct for the fact that variance penalizes over- and underperformance equally, are always attained in a single period irrespective of the market condition or the security return distribution. Leland (1999), in turn, provides an approach beyond mean-variance analysis: He argues that a simple modification of the CAPM beta can produce correct risk measurement for portfolios with arbitrary return distributions, and the resulting alphas of fairly priced options, as well as dynamic strategies, will be zero. This risk measure requires no more information to implement than the CAPM.

Furthermore, recall that in the Markowitzian portfolio theory, the portfolio allocation depends upon the investor's utility function. Stigler (1965) discusses the initial development of the utility function, whereas Markowitz and van Dijk (2006, 180–184) and Infanger (2006, 209–214) illustrate further the use of various utility functions in the mean-variance framework. Indifference curves, risk aversion, and investor risk preferences are also addressed in Sharpe, Alexander, and Bailey (1999). As noted by Wilmott (1998, 538), the approach of choosing the optimal portfolio with the aid of a utility function is particularly popular with economists. Economists also generally agree that a reasonable utility function should display non-satiety, risk aversion, decreasing absolute risk aversion, and constant relative risk aversion. It can be shown that the only utility functions that satisfy all these desirable characteristics are the logarithmic function ($\log W$) and the power function ($W^{1-\gamma}$). (see Vander Weide 2010, 162.)

In practice, however, it can be challenging, or even impossible, to determine an investor's actual utility function. For this reason, Roy (1952) argues that an investor, rather than thinking in terms of utility functions, first wants to ensure that a certain amount of the principal is preserved. After that, she decides on some minimal acceptable return that achieves this principal preservation. This 'safety-first' approach can be seen to have laid the seed for the development of downside risk measures, such as lower partial moment (see, e.g., Bawa 1978; Nawrocki 1999; Fabozzi et al. 2010, 345–346). Correspondingly, Kahneman and Tversky (1979) and Kahneman and Riepe (1998) consider investor psychology and focus on studying loss aversion. Chelo (2000), in turn, demonstrates how spreads can be used to profit from these kinds of behavioral 'biases.'

In contrast to the classical mean-variance analysis, Luenberger (1998, Chapter 15) considers the theory of 'optimal portfolio growth' and 'log-optimal portfolios.' This long-run growth rate maximization and the optimal investment amount is generally known as the 'Kelly criterion' (see Kelly 1956). In turn, Baz and Guo (2017) describe and contrast the mechanics of standard asset allocation models, including the utility-based, Markowitz, Kelly, risk parity, and fixed allocation approaches. They also identify the precise conditions under which the models are equivalent. Recently, growth optimal investment strategies have also been combined with state-of-the-art machine learning algorithms. For more information on such approaches, see, e.g., Horváth and Urban (2010), Györfi, Ottucsák, and Walk (2012), Li and Hoi (2012; 2014; 2016), Zhang et al. (2020), as well as Zhang (2021). For example, the online portfolio selection strategies proposed by Li and Hoi (2016) maximize the expected log return of a portfolio and aim for multiple-period portfolio selection. For alternatives to the Kelly criterion, see, e.g., Browne (1999; 2000). He finds a dynamic strategy that maximizes the probability of reaching a given wealth level in a specified time and shows that this sizing strategy is equivalent to the hedging strategy of a binary call (see also Sinclair 2013, 150–153 for a simple demonstration).

2.2 Problems with the classical mean-variance optimization

Despite the theoretical simplicity and the intuitive appeal of the modern portfolio theory, fundamental challenges are commonly encountered when applying it to real-world portfolio construction. Indeed, the first issue with the mean-variance framework is that it requires numerous parameters to be estimated (i.e., n for the expected returns of the assets and $n(n+1)/2$ for their covariances). Second, mean-variance optimization is generally very sensitive to its inputs, and minor deviations often lead to significant changes in portfolio weights. Essentially, this sensitivity problem is related to the fact that the mean-variance optimizer does not know that the inputs are statistical estimates and not known with certainty. In other words, it is implicitly assumed that inputs are deterministic and available with great accuracy. (Fabozzi et al. 2010, 337.)

Furthermore, observations of the historical asset return data are needed to estimate the expected return and the covariance matrix by the sample mean and sample covariance matrix. Unfortunately, such a historical approach may lead to a counterintuitive or unstable outcome because statistical estimates are noisy and depend on the quality of the data and the statistical techniques used. One may, for example, need to deal with missing or truncated data (see Little & Rubin 2002; Stambaugh 1997, respectively). Indirectly, it is also assumed that the past historical performance can predict the future, even though it is well known that markets and economic conditions change throughout time. In general, the estimation error is worsened for both the expected return and covariance estimators when the number of historical return observations is small relative to the number of securities. Respectively, it can be shown that only in the case of independent and identically distributed (i.i.d.) time series are the sample mean and covariance estimator the maximum likelihood estimators of the true mean and covariance (see, e.g., Hayashi 2000). Hence, the forecasting power of sample mean and covariance estimators (as a forecast of expected return and risk) is typically poor, and modifications and extensions may be necessary for practical applications. (Fabozzi et al. 2010, 331, 334–336, 341.)

Overall, several conclusions have been made in the literature regarding the above-mentioned issues. Firstly, uncertainty from estimation error in expected returns tends to have more influence than in the covariance matrix in the mean-variance allocation process (see, e.g., Best & Grauer 1991; 1992). Indeed, although the relative importance depends on the investor's risk aversion, errors in the expected returns are generally about ten times more critical than errors in the covariance matrix, whereas errors in the variances are about twice as important as errors in the covariances (Chopra & Ziemba 1993, 7). As the risk tolerance increases, the relative impact of estimation errors in the expected returns becomes even more important. Therefore, the primary focus should be on providing good estimates for the expected returns, followed by the variances. Due to the 'error maximization' property, MV optimization significantly overweights (underweights) those securities that have large (small) estimated returns, negative (positive) correlations, and small (large) variances (Michaud 1989, 33–34). Hence, mean-variance optimized portfolios are not necessarily well-diversified since they often consist of a few large and many small positions (Jorion 1985, 261). It has also been documented that equally-weighted portfolios often outperform mean-variance optimized portfolios (see, e.g., Jobson & Korkie 1981; DeMiguel, Garlappi & Uppal 2009). (see also Fabozzi et al. 2010, 365.)

To some extent, the issue related to parameter uncertainty can be mitigated by using robust or stable estimators of location (μ) and scatter (σ) (see, e.g., Huber 1981; Welsch & Zhou 2007; Maronna et al. 2019). Another popular approach has been to generate expected return forecasts using different factor models based on Arbitrage Pricing Theory (Ross 1976) (see, e.g., Duffie 1988; Luenberger 1998). Naturally, the utilization of factor models also helps to reduce the dimensionality of the portfolio optimization problem (see, e.g., Chan, Karceski & Lakonishok 1999). Aguilar and West (2000), in turn, suggest forecasting the covariance matrix using dynamic factor models, which are direct generalizations of univariate stochastic volatility models, as opposed to generalized ARCH structure (cf., e.g., Demos & Sentana 1998; Engle & Sheppard 2001). Finally, Elton et al. (2014) demonstrate that, in the case of single-factor models, the resulting simple covariance matrix leads to very efficient portfolio selection algorithms.

An alternative approach to mitigate the problems of the classical mean-variance allocation process has been to utilize Bayesian or shrinkage estimators (see, e.g., Jorion 1986; Black & Litterman 1992; Kandel & Stambaugh 1996; Pastor 2000; Pastor & Stambaugh 2000; Ledoit & Wolf 2003; 2004; Garlappi, Uppal & Wang 2006). As noted in Black and Litterman (1992, 28, 30, 33), investors using the standard model must translate their views into a complete set of expected excess returns on assets that can be used as a basis for portfolio optimization. In addition to the practical estimation difficulties, this model cannot distinguish strongly held views from auxiliary assumptions, and the optimal portfolio it generates often appears to bear little or no relation to the views the investor wishes to express.

On the other hand, a Bayesian framework can be utilized to incorporate the investor's prior degree of confidence, e.g., in an asset pricing model. For example, in the model proposed by Kandel and Stambaugh (1996, 349), a Bayesian investor uses the sample evidence to update prior beliefs about the regression parameters and then uses these revised beliefs (Bayesian posterior distributions) to compute the optimal asset allocation. Similarly, Pastor (2000, 181) proposes an approach that shrinks the sample mean toward their values implied by the CAPM. The extent of deviations from the model depends on the strength of the violations of the model in the data and on the investor's degree of confidence in the model. Shrinking the sample mean tends to reduce the sensitivity of the optimal weights to the sampling error. Thus, the weights have less extreme values and less variability than in the data-based approach (see, e.g., Larsen & Resnick 2001).

Several other shrinkage (i.e., averaging) approaches have also been suggested. For instance, Jorion (1986) develops a so-called 'Bayes-Stein' estimator for the expected return, which shrinks toward the return of the global minimum variance portfolio. Ledoit and Wolf (2003; 2004), in turn, analyze optimal shrinkage intensity and utilize a covariance matrix that follows from the single-factor model developed by Sharpe (1963) or the constant correlation model (cf. Elton & Gruber 1973). The authors observe that there is about four times as much estimation error present in the sample covariance matrix as there is bias in the single-factor covariance matrix. The results of Ledoit and Wolf (2003; 2004) demonstrate that when it comes to computing a GMV portfolio, shrinkage estimators are superior compared to, e.g., statistical and fundamental factor models. (see also Martellini & Ziemann 2008, 3, 6, 10; Fabozzi et al. 2010, 337–342, 371–373.)

However, it is worth noting that an opposite conclusion is made in Disatnik and Benninga (2007, 62): They find, quite surprisingly, that while portfolios of estimators (see, e.g., Bengtsson & Holst 2002), as well as shrinkage estimators of the covariance matrix, are undeniably better than the simple sample covariance matrix estimator, there are no statistically significant differences in portfolio performance over time between stock portfolios constructed using simple portfolios of covariance matrix estimators and stock portfolios constructed using more sophisticated shrinkage estimators of the covariance matrix, at least when monthly data is used, and no short-sale constraints are imposed.

Based on these previous papers, Garlappi et al. (2006) examine the normative implications of parameter and model uncertainty for asset allocation, using a model that allows for multiple priors and where the decision-maker is averse to ambiguity. The authors characterize the multiple priors by a 'confidence interval' around the estimated expected returns and model ambiguity aversion via a minimization over the priors. Their empirical analysis suggests that, compared with portfolios from classical and Bayesian models, ambiguity-averse portfolios are more stable over time and deliver a higher out-of-sample Sharpe ratio.

Interestingly, Kan and Zhou (2007) find that with parameter uncertainty, holding the two-fund portfolio consisting of the sample tangency portfolio and the riskless asset is never optimal, and an investor can benefit by holding some other risky portfolios that help reduce the estimation risk. More specifically, the authors show that a so-called three-fund portfolio, which optimally combines the riskless asset, the sample tangency portfolio, and the sample global minimum-variance portfolio, dominates the two-fund portfolio. In contrast, Brandt, Santa-Clara, and Valkanov (2009) propose a methodology where parametric portfolios are chosen by modeling the portfolio weight directly in each asset as a function of the stock-return characteristics, such as book-to-market ratio, size, and momentum. This computationally simple approach appears to produce sensible portfolio weights and offer robust performance in- and out-of-sample, as opposed to the traditional Markowitz approach. Finally, Filomena and Lejeune (2012, 212) address the issue of estimation risk by arguing that investors would rather trade off some return for more safety and construct a portfolio that performs well under a wide set of circumstances. For this reason, the authors focus on stochastic optimization techniques and propose a probabilistic version of the Markowitz model in which the incomplete knowledge of the return behavior is taken into account by modeling the asset returns as random variables.

Studies have also shown that the inclusion of constraints in the mean-variance optimization problem may improve risk-adjusted out-of-sample performance (see, e.g., Frost & Savarino 1988; Grauer & Shen 2000; Clarke, de Silva & Thorley 2002). Practitioners often use no short-selling constraints or upper and lower bounds for each security to avoid overconcentration in a few assets. Jagannathan and Ma (2003) illustrate further that the no short-selling constraints are equivalent to reducing the estimated asset covariances, whereas upper bounds are equivalent to increasing the corresponding covariances. Moreover, solving the global minimum variance portfolio problem with some constraints on weights is equivalent to using a shrinkage estimate of the covariance matrix. However, if the constraints used for stability and robustness purposes are too tight, they will completely determine the portfolio allocation and not the forecasts. (Fabozzi et al. 2010, 365–366; see also Roncalli 2010 for more discussion on the impact of weight constraints in portfolio theory.) Recently, there have also been investigations on the performance of portfolios generated by stochastic programming with multivariate second-order stochastic dominance constraints (see Meskarian, Fliege & Xu 2014). These optimization problems are essential in multi-criterion decision-making since each component of vectors can be interpreted as the uncertain outcome of a given criterion. Promisingly, the results by Meskarian et al. (2014, 113) suggest that their portfolio strategy significantly outperforms the portfolio generated by the Markowitz model.

Data frequency (or the length of the estimation period) is another crucial part of the accuracy of mean-variance optimization. Merton (1980) shows that even if the expected returns were stationary (i.e., constant over time), an exceedingly long history would still be required to estimate them accurately. It is also well-known that the sample mean is the Best Linear Unbiased Estimator (BLUE) of the population mean for distributions that are not heavy-tailed. In such a case, the sample mean exhibits a property stating that an increase in the sample size always improves its performance. (Fabozzi et al. 2010, 336.) However, these results are no longer valid under extreme thick-tailedness (see, e.g., Ibragimov 2007).

The situation is quite different for variances and covariances: Butler and Schachter (1986) point out that when historical data is used for volatility forecasting purposes, the bias found in the estimator tends to increase with the sample length. On the other hand, it can be problematic to use information based on too short time periods because, in such a situation, the volatility estimator often becomes highly sensitive to short-term regimes, such as over- and underreaction corrections. Although an increase in the sampling frequency may improve estimates of these quantities under reasonable assumptions, such high-frequency or tick-by-tick data needed is rarely available for free. (Fabozzi et al. 2010, 341–342.) Therefore, as discussed in Garman and Klass (1980), a natural compromise could be to employ data based on the historical opening, closing, high and low prices, and transaction volume. On the other hand, authors such as Litterman and Winkelmann (1998), De Santis et al. (2003), and Pafka, Potters, and Kondor (2004) argue that weighted data should be used to account for the fact that as the market changes, more importance should be given to more recent information. In any case, it seems reasonable to test the estimator of the covariance matrix (or the expected return) for the specific asset classes and dataset with which a practitioner is dealing before adopting it for portfolio management purposes (Fabozzi et al. 2010, 340). For example, autocorrelation and heteroskedasticity, which are commonly found in financial return series (see, e.g., Campbell, Lo, and MacKinlay 1997), introduce biases in the estimated covariance matrix. Fortunately, techniques such as 'Newey-West corrections' (Newey & West 1987) can correct these biases even in the portfolio construction context (see Rab & Warnung 2011).

Instead of using improved estimators, estimation errors can also be incorporated directly into the optimization process by utilizing robust portfolio optimization techniques (see, e.g., Ben-Tal & Nemirovski 1998; Tütüncü & Koenig 2004; Ceria & Stubbs 2006; Todorov & Filzmoser 2009; Scherer 2010; Cornuéjols, Peña & Tütüncü 2018 for detailed expositions). In brief, the uncertain parameters in the optimization problem are assumed to vary in prespecified uncertainty sets that are selected based on statistical techniques and probabilistic guarantees. Under certain conditions, it can be further shown that the optimal portfolio weights using robust optimization are a linear combination of the weights of the minimum variance portfolio and a mean-variance efficient portfolio with speculative demand. Thus the implied expected return is equivalent to the expected return obtained using a shrinkage estimator with certain weights. (see Scherer 2007 for the proof.) As noted by Fabozzi et al. (2010, 395, 411–412, 418), investors using robust portfolio optimization formulations are likely to trade off the optimality of their portfolio allocation in cases where nature behaves as they predicted for protection against the risk of inaccurate estimation. Thus, investors using such techniques should not expect to outperform classical optimization when estimation errors have little impact or when typical scenarios occur. However, they should expect insurance in scenarios where their estimates deviate from the actual realized values to the amount they have prespecified in the modeling process.

The robustification framework has been widely studied in the literature. For example, Goldfarb and Iyengar (2003) propose a robust portfolio selection model where the mean return vector and the covariance matrix (i.e., the market parameters) lie in known and bounded uncertainty sets. The authors then compute a portfolio by solving a max-min mean-variance problem assuming the worst-case behavior of the parameters. Erdogan, Goldfarb, and Iyengar (2003) demonstrate further how robust optimization-based techniques can be utilized to immunize the active portfolio selection problems to perturbations in parameter values. Bonami and Lejeune (2009), in turn, show that robust constraints can extend the standard MV portfolio optimization model. In their exact solution approach, integer trading restrictions and the uncertainty in the estimate of the expected returns are simultaneously considered (see also Uryasev 2013; Drewes 2016 for practical applications based on this model). Finally, Shaw (2011) illustrates that robustification is possible even when using Monte Carlo sampling of random portfolios to solve investment problems.

All these various approaches can be, to some extent, supported by empirical findings. For example, Ceria and Stubbs (2006, 126) observe that portfolios constructed using robust optimization typically outperform those created using traditional mean-variance optimization in terms of annualized returns. Similarly, Tütüncü and Koenig (2004, 183) find that the worst-case behavior of portfolios of different asset classes can be improved significantly using robust asset allocation methods. Robust efficient portfolios also remain relatively unchanged and stable over long periods of time, although they may be concentrated on a small set of asset classes. On the other hand, Scherer (2007, 386) argues that robust optimization underperforms traditional MV optimization in terms of expected utility if investors show little risk aversion but high uncertainty aversion. Finally, El Ghaoui, Oks, and Oustry (2003, 555) consider the problem of optimizing the worst-case Value-at-Risk when the distribution of returns is partially known in the sense that only bounds on the mean and covariance matrix are available. They note that data errors dramatically impact the VaR if the nominal portfolio (i.e., the one resulting from the assumption that the data are error-free) is chosen instead of an optimally robust portfolio.

Since expected returns can rarely be estimated with enough accuracy, several authors have chosen to drop them altogether and focus only on the covariance matrix. This has led to risk-based asset allocation approaches, of which ‘risk parity’ is a prominent example. By definition, the risk parity approach aims to construct portfolios where the overall risk is diversified by allocating the risk equally across different securities or investment strategies. The portfolio risk and the risk contributions are typically calculated from the variance and covariance estimates of their future returns. As argued in Kolm, Tütüncü, and Fabozzi (2014, 365), risk parity portfolios have appealing properties that make them appropriate choices for investors looking for diversified portfolios. For example, Maillard, Roncalli, and Teïletche (2010) have shown that a risk parity portfolio’s volatility is located between those of minimum variance and equally-weighted portfolios and that it is a good trade-off between those approaches in terms of the absolute level of risk, risk budgeting, as well as diversification. Furthermore, Asness, Frazzini, and Pedersen (2012) find that risk parity portfolios overweight safer assets relative to their weights in the market portfolio and that risk parity portfolios can be leveraged to achieve the same risk as to the market portfolio and a higher expected return. Anderson, Bianchi, and Goldberg (2012) continue by noting that although unlevered risk parity strategies have high Sharpe ratios and their performance improves in turbulent periods, a risk parity strategy levered to match market volatility underperforms the market based on both return and risk-adjusted return after accounting for borrowing and trading costs. (see also Kaya & Lee 2012; Jurczenko 2015 for more discussions.)

Unfortunately, dropping the forecasts on returns improves but does not prevent the instability issues. This is because quadratic programming methods require the inversion of a positive-definite covariance matrix. It can be shown that this inversion is prone to large errors when the covariance matrix is numerically ill-conditioned, i.e., when it has a high condition number. The condition number rises quickly if correlated (multicollinear) assets are added, and, as the size of the covariance matrix increases, each covariance coefficient is estimated with fewer degrees of freedom. Hence, at some point, the benefits of diversification are more than offset by estimation errors. (López de Prado 2018, 222–223; see also Bailey & López de Prado 2012.)

In order to address these concerns, López de Prado (2016; 2018) proposes a novel approach called Hierarchical Risk Parity (HRP), which applies graph theory and machine learning-based tree structures to build a diversified portfolio based on the information contained in the covariance matrix. Unlike quadratic optimizers, HRP does not require the invertibility of the covariance matrix, meaning that it can compute a portfolio on an ill-degenerated or even a singular covariance matrix. The author's Monte Carlo experiments show that HRP delivers lower out-of-sample variance than the classical Markowitzian optimization approach and produces less risky portfolios compared to traditional risk parity methods. (see López de Prado 2018, 221–223, 236–238.) Further support for these claims is provided in Raffinot (2018), who finds that HRP portfolios yield attractive risk-adjusted returns and match with volatility targets much better than minimum variance portfolios. Subsequently, Jain and Jain (2019) discover that HRP is less sensitive to covariance misspecification when compared with minimum variance or maximum diversification portfolio, while it is not as robust as the inverse volatility weighted portfolio.

Still another fascinating perspective to investment management is presented in Martellini and Ziemann (2008, 3). They argue that, as most asset returns exhibit deviations from normality, optimal portfolio selection techniques should involve higher-order comoments of asset return distribution as inputs, in addition to the covariance matrix. It has also been documented that investors typically exhibit non-trivial preferences with respect to portfolio higher-order moments, i.e., skewness and kurtosis (see Scott & Horvath 1980). However, given the enormous increase in dimensionality involved in higher-order moment parameter estimation, it is challenging to efficiently implement such an alleged improvement over the classical mean-variance portfolio selection in practical situations. Nevertheless, various proposals have been made. For example, Harvey et al. (2010) follow a Bayesian decision-theoretic framework that addresses the ability to handle higher moments and parameter uncertainty. A comparison to other methods where parameter uncertainty is either ignored or accommodated in an *ad hoc* way shows that such an approach can lead to higher expected utility than, e.g., the resampling methods (see Michaud 1998). On the other hand, the approach proposed by Harvey et al. (2010, 482) has several limitations. Most notably, the authors' framework is applied only to in-sample portfolio selection, and the examined portfolio choice problem is a static one, as opposed to a dynamic asset allocation problem with varying portfolio weights. Note that we return to this issue in Subchapter 3.3, where we show that the impact of higher-order moments of asset returns on portfolio selection techniques can alternatively be assessed by utilizing a statistical factor model.

Finally, it is worth mentioning that there are also multi-period extensions of mean-variance optimization. Unlike single-period portfolio choice policies, those can incorporate intertemporal effects such as changing market conditions, market impact costs, hedging needs, and alpha decay. As discussed in Kolm et al. (2014, 368–369), return predictability and market impact give rise to intertemporal hedging demands for securities, and investors need to look beyond just the next period when optimally allocating across assets. For example, market impact costs from trades in the current period may affect prices in later periods. Without going into technical details, it can be shown that since the multi-period portfolio optimization problem is typically formulated as a Linear-Quadratic-Gaussian (LQG) control problem, the optimal trade is proportional to the difference between current portfolio holdings and dynamically modified optimal target portfolio. Hence, portfolios are moving targets that are optimally tracked to balance the trade-off between risk-adjusted returns and trading costs while accounting for the persistence effects of both. However, there are several reasons why practitioners typically resort to single-period models to rebalance the portfolio from one period to another. For instance, it is well-established that accurate estimation of returns and risk is exceedingly tricky, especially for multiple periods. Moreover, the most common existing multi-period models do not handle real-world constraints. An interesting exception here is provided by Kolm and Ritter (2015): They present a new theoretical framework for multiperiod optimization with transaction costs and constraints, which recasts the problem as an estimation of a hidden state sequence in a Markov chain. (see also Garleanu & Pedersen 2013; Boyd et al. 2014 for more discussion on multiperiod portfolio selection.)

2.3 Black-Litterman model and its variations

In this subchapter, we focus more closely on the Black-Litterman (BL) model and its variants because, as we see in Subchapters 3.4–3.5, it can provide a versatile framework for constructing tactical asset allocation strategies exploiting, e.g., different anomalies or time-series forecasts. Moreover, as argued by Idzorek (2007, 37), the BL model can overcome some of the well-known weaknesses of mean-variance optimization, namely unintuitive and highly concentrated portfolios, input-sensitivity, and estimation error maximization.

During the thirty or so years since the original papers by Black and Litterman (1991ab; 1992), many authors have published research referring to their model as Black-Litterman, even though these models' formulations may not be based on the same underlying principles as the 'canonical model.' According to Walters (2014, 2), this canonical model makes two notable contributions to the problem of portfolio selection: First, to improve the estimation of asset returns and shrinkage, 'reverse optimization' can be used to generate a stable distribution of returns from the prior equilibrium market portfolio. Second, the canonical model provides a quantitative framework for specifying the investor's relative or absolute views on some or all of the expected returns and a way to combine those views with prior information to arrive at a new combined distribution. The investor can also control how strongly a view influences portfolio weights, in conformity with the degree of confidence with which she holds the view. Finally, it is assumed that Gaussian probability distribution is suitable for the means of the random components. (see Black & Litterman 1992, 1–2, 28, 33–35.)

Mathematically, the Black-Litterman expected returns formula is a confidence-weighted linear combination of market equilibrium and the expected return implied by the investor's views, and it can be expressed as follows:

$$\hat{\boldsymbol{\mu}}_{BL} = [(\boldsymbol{\tau\Sigma})^{-1} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}]^{-1}(\boldsymbol{\tau\Sigma})^{-1}\boldsymbol{\Pi} + [(\boldsymbol{\tau\Sigma})^{-1} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}]^{-1}[\mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}]\boldsymbol{\mu}, \quad (2.3.1)$$

or, equivalently,

$$\hat{\boldsymbol{\mu}}_{BL} = [(\boldsymbol{\tau\Sigma})^{-1} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}]^{-1}[(\boldsymbol{\tau\Sigma})^{-1}\boldsymbol{\Pi} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{q}], \quad (2.3.2)$$

where

$\boldsymbol{\Sigma}$ is the $N \times N$ covariance matrix of returns.

$\boldsymbol{\Pi}$ is $[\Pi_1, \dots, \Pi_N]'$ and stands for the vector of expected (excess) returns. It can be computed from historical returns or from an equilibrium model such as the CAPM or a multifactor pricing model based on the APT. Idzorek (2007, 19-20) shows that this vector of implied excess (or mean) equilibrium returns can be extracted from known information. In other words, $\boldsymbol{\Pi} = \boldsymbol{\lambda\Sigma w}_{mkt}$, where $\boldsymbol{\lambda}$ is the risk-aversion coefficient, and $\boldsymbol{w}_{mkt}$ is the market capitalization weight ($N \times 1$ column vector) of the assets. The scaling factor, $\boldsymbol{\lambda}$, can be estimated by dividing the expected excess return by the portfolio variance, and thus the weighted reverse optimized excess returns equal the specified market risk premium (see also Grinold & Kahn 2000).

$\boldsymbol{\tau}$ is a scalar representing the confidence in the estimation of the market prior. In other words, it measures the uncertainty of the equilibrium variance.

$\mathbf{q}$ is a K-dimensional vector of K investor views (i.e., a vector of the estimated excess returns for each view).

$\mathbf{P}$ is a $K \times N$ 'pick' matrix of the asset weights within each view. For a relative view, the sum of the weights will be zero, and for an absolute view, the sum of the weights will be one. Different methods have been proposed to compute these weights: For example, He and Litterman (2002) and Idzorek (2007) use a market capitalization-weighted scheme, whereas Satchell and Scowcroft (2000) use an equal-weighted scheme. Litterman (2003, 82) assigns a percentage value to the assets in question. In practice, weights will be a mixture depending on the process used to estimate the view returns (Chan 2017, 137).

$\boldsymbol{\Omega}$ is a $K \times K$ matrix of the covariance of the views (i.e., it expresses the confidence in the investor's views). $\boldsymbol{\Omega}$ is often assumed to be diagonal, meaning that investor views are assumed to be independent (see, e.g., He & Litterman 2002, 4). However, alternative approaches have been proposed for calibrating $\boldsymbol{\Omega}$ (Allaj 2013, 224): For example, Meucci (2007, 430–431) writes this uncertainty matrix as $\boldsymbol{\Omega} \equiv \left(\frac{1}{c} - 1\right) \mathbf{P\Sigma P}'$, where c is a positive scalar. In his parameterization, the scalar c tweaks absolute confidence in the investor's skills. The case $c \to 0$ gives rise to an infinitely dispersed distribution of the views (i.e., the investor's views have no impact). The case $c \to 1$ gives rise to an infinitely peaked distribution of the views (i.e., the investor is entirely trusted over the official market model). Thus, the case $c \equiv 1/2$ corresponds to the situation where the investor is trusted as much as the official market model.

The covariance of the Black-Litterman estimator of expected returns is given as:

$$[(\boldsymbol{\tau\Sigma})^{-1} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}]^{-1}, \tag{2.3.3}$$

and it can be used as an approximation for the estimation error covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\mu}}$. It is worth noting that since security returns are correlated, views only on a few assets will imply changes to the expected returns on all assets. Any estimation errors are also spread out over all assets, which makes the BL expected return vector less sensitive to errors in individual views. Such an effect contributes to the mitigation of estimation risk in the optimization process. (Fabozzi et al. 2010, 375–382, 404, 412; see also Chan 2017, 143–144, 136–137, 182 for the complete derivation of the BL 'master formula.')

Unfortunately, the original formulation of the Black-Litterman model raises a typical practical issue, which is the curse of dimensionality. This is because the posterior covariance matrix produced by the BL model is a fully populated numerical matrix. In order to reduce the dimensions, Cheung (2010, 237–239) recommends an analysis based on sorted eigenvalues (i.e., an eigensystem analysis). This solution obtains a sparse matrix representation by utilizing the principal component approximation. Alternatively, it would be possible to rank the securities according to the posterior estimates and reduce the universe before optimization. On the plus side, the normality and linearity assumptions of the BL model allow an analytical solution to the optimization problem, meaning that the implementation can be efficient. Moreover, since the outputs are in the form of explicit return forecasts together with a covariance matrix, a standard mean-variance optimizer can produce robust allocations due to the shrinkage effect.

Another well-known practical issue is the correct setting of the confidence parameter $\boldsymbol{\tau}$, as there seems to be no consensus of its calibration in the literature (see, e.g., Allaj 2013, 219, 229; Walters 2014, 22–25). For example, Chan (2017, 150) argues that since $\boldsymbol{\tau\Sigma}$ is analogous to the standard error, one should set *tau* inversely proportional to the number of the available observations, i.e., $\boldsymbol{\tau} = 1/T$ (the maximum-likelihood estimator), or $\boldsymbol{\tau} = 1/(T - k)$ (the best quadratic unbiased estimator), where T is the number of samples and k is the number of assets. According to Walters (2014, 24), this technique is consistent with several papers indicating that the used values of $\boldsymbol{\tau}$ are small and typically on the range of $(0.025, 0.05)$ (see, e.g., Black & Litterman 1992; He & Litterman 2002; Idzorek 2007). Beach and Orlov (2007, 162) continue by showing that the highest portfolio variance and the most extreme allocations occur when $\boldsymbol{\tau}$ is set to 0.01 or higher.

It is worth noticing that, when the so-called Alternative Reference Model is used, the meaning and impact of the parameter $\boldsymbol{\tau}$ deviate from the Canonical Reference Model (see, e.g., Walters 2014, 3–4, 20–25; Chan 2017, 143, 146–150, 189). Mathematically, the definition of the expected returns changes from $\mathrm{E}(\boldsymbol{r}) \sim \mathcal{N}(\boldsymbol{\Pi}, (1 + \boldsymbol{\tau})\boldsymbol{\Sigma})$ to $\mathrm{E}(\mathbf{r}) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. This means that the Alternative Reference Model does not consider the market mean a random variable and does not include updating the covariance of the estimates. The first example of such a model is provided by Satchell and Scowcroft (2000, 138–141): In their analysis, point estimates are used, and $\boldsymbol{\tau}$, which now acts as a known scaling factor, is set to one. Therefore, the BL model is treated as a shrinkage or mixing model. Satchell and Scowcroft (2000, 144–145) also formulate $\boldsymbol{\tau}$ as an unknown and stochastic parameter, but this method is challenging to implement due to the additional parameters introduced. The second example is presented by Meucci (2008a, 7–8): In his modification, $\boldsymbol{\tau}$ is not included, and views are formed directly on market realization rather than the mean (see also Gochez 2015; Pfaff 2016, 291).

Further insights into an appropriate $\boldsymbol{\tau}$ value-setting are also provided by Allaj (2013): He develops an econometric estimation procedure that captures the serial, cross, and cross-lag correlations of the excess returns and yields a suitable and consistent point estimate of $\boldsymbol{\tau}$. The estimated parameter $\hat{\boldsymbol{\tau}}$ is also allowed to take any value, not only values in $[0,1]$. Although the posterior expected excess returns obtained with this reformulated BL model can be significantly different from those obtained with the original BL model, Allaj (2013, 238, 244–246) argues that the utilization of such an approach may lead to more diversified portfolios. Tee, Huang, and Lim (2017, 126–128, 130–131), in turn, show that the BL model's uncertainty factor and idiosyncratic risk aversion, which are arbitrarily fixed in most applications, are related to the Sharpe ratio and Value-at-Risk (i.e., tail risk) of the active portfolio. The other observation made by the authors is that these parameters are not entirely exogenous but are connected closely to the investor's inputs of subjective expected returns and the degree of confidence over these beliefs. Hence, the BL model can become internally consistent if the parameter calibrations are unified with performance and risk measures.

Overall, the original BL model and its several modifications and extensions have been widely applied in the literature. For instance, it has been demonstrated that the model can incorporate other asset classes beyond equities and bonds (see, e.g., Black & Litterman 1991; Bevan & Winkelman 1998). He and Litterman (2002), in turn, measure the impact of each view on the final posterior weights, whereas Fusai and Meucci (2003) quantify the statistical difference between the posterior return estimates and the prior estimates. Pézier (2007) processes full and partial views on expectations and correlations by least discrimination, while Idzorek (2007) presents a method for assessing the variance of the views. Braga and Natale (2007) calibrate the uncertainty in views by utilizing a measure for the marginal contribution of each view to the ex-ante Tracking Error Volatility (TEV). In contrast, da Silva, Lee, and Pornrojnangkool (2009) apply the BL model to active portfolio management by maximizing the portfolio's total return with a penalty on the tracking error.

Furthermore, Fabozzi et al. (2010, 384, 394) show that factor models and simple ranking models can be simultaneously incorporated into the Black-Litterman model. Similarly, Krishnan and Mains (2005) include in the BL model factors that are not priced in the market (such as a recession indicator), whereas Figelman (2017) applies the BL framework with a factor structure to multi-asset portfolios (see also Jones, Lim & Zangari 2007). Almgren and Chriss (2004, 2005), in turn, express 'lax' views (e.g., ranking) on expectations in a framework similar to the BL model. Finally, Babameto and Harris (2008) develop forecasting models for zero-investment value and momentum strategies and incorporate the predictions from these models into the BL approach. They also show that such an approach can outperform the benchmark market index, even in the presence of substantial transaction costs.

With respect to risk modeling, Giacometti et al. (2007) employ more realistic models for asset returns (such as Student's t) and extend the BL model to the stable distribution family. The authors also solve the portfolio allocation problem for various risk measures, including dispersion-based risk measures, Value-at-Risk (VaR), and Conditional-Value-at-Risk (CVaR). Martellini and Ziemann (2007) apply the BL model to portfolio construction in the presence of non-trivial preferences about higher moments of asset return distributions and use VaR as the objective function. Lejeune (2011) proposes a Value-at-Risk Black-Litterman model that accounts for the VaR and various trading constraints. He then uses this approach to construct fund-of-funds that embed an absolute return strategy.

Much attention has also been paid to incorporating volatility forecasts into the BL model. For example, Satchel and Scowcroft (2000) derive a model where an investor can present views on global volatility, while Qian and Gorman (2001) propose a similar approach based on conditional distribution theory. In the setting used by Beach and Orlov (2007, 148–149), an exponential GARCH-in-Mean model is employed to produce estimates on expected returns and variances. The authors argue that the returns of their portfolio were able to surpass those of portfolios that rely on market equilibrium weights or Markowitz-optimal allocations because the asymmetric GARCH models can systematically capture many regularities of stock returns. These findings are in line with Duqi, Franci, and Torluccio (2014), who also utilize the EGARCH-M model to generate views, and observe that greater reliance on the implied BL excess returns improves the risk-return ratio.

Furthermore, Palomba (2008) provides an empirical model for large-scale tactical asset allocation with multivariate GARCH estimates, given a tracking error constraint. In his paper, the BL approach is utilized to manage the selected portfolio by combining information taken from the time-varying volatility model with personal views about asset returns. Wang (2010, 391), in turn, incorporates the BL model with Markov regime-switching in GARCH processes (MRS-GARCH) so that the forecasts from the MRS-GARCH model represent the investor’s views. This allows him to study the impact of volatility shifts between two regimes on a dynamic asset allocation strategy. Ammann and Verhofen (2006) demonstrate further that high-variance and low-variance regimes exhibit significantly different optimal allocations both in value and momentum stocks, implying that allocation decisions based on variance can be considered beneficial.

On a different note, Rebonato and Denev (2013) highlight the possible instability issues of the Black-Litterman framework caused by Bayesian inference. More specifically, the authors present subjective views about mean asset returns as well as co-dependencies and use ‘Bayesian nets’ as tools since they can deal consistently and coherently with scenario analysis and stress events. Rebonato and Denev (2013) then compare the original BL solution to a geometric version of mean-variance optimization and show that by utilizing the latter approach, it is possible to have stable and well-behaved weights while remaining closer than the BL model to the investor’s views. They also find that the inclusion of low-probability tail events plays an important role in portfolio optimization. (see Rebonato & Denev 2013, 5, 76, 83, 90, 465 for more details.)

More criticism against the Black-Litterman model has been presented by Michaud, Esch, and Michaud (2012). They show that constrained BL is identical to Markowitz and that portfolios based on Monte Carlo resampling methods are better diversified and risk-managed under identical inputs and optimality criteria. Correspondingly, Meucci (2005, 2) notes that the BL approach suffers at least from two drawbacks. Firstly, the assumption that both the market prior and the investor's views are normally distributed is too strong in most contexts. Fortunately, this normality assumption can be overcome within the BL framework if one accepts numerical results instead of closed analytical formulas. The second problem is related to the fact that the investor is supposed to express views on the parameters determining the market distribution instead of the possible realizations of the market. The dichotomy between the subjective views and the prior market distribution can be solved by utilizing the so-called 'opinion pooling' approach, which extends the BL methodology to generic non-normal market distributions and non-normal views. It also allows one to determine the marginal distribution of each view separately, and the joint co-dependence (i.e., the copula) is directly inherited from the prior market structure among the views. With a suitable change of coordinates, the joint distribution of the views can then be translated into a joint posterior distribution for the market. (Meucci 2005, 2–5; see also Genest & Zidek 1986; Clemen & Winkler 1999; Meucci 2006 for more discussion on opinion pooling methods.)

Meucci (2005, 11–14) empirically examines his model in a fat-tailed, skewed, and highly dependent market, where the views are expressed as uncertainty ranges. He finds that the opinion pooling posterior represents a gentler modification of the prior than the BL posterior, resulting in a less pronounced twisting of the optimal allocations in a portfolio optimization context. However, unlike in the BL model, where the posterior distribution is again normal, in the opinion-pooling approach, the posterior can be highly non-symmetrical. Hence, asymmetries and tail risk can come to play a significant role, and the mean-variance optimization technique can become highly sub-optimal. To overcome this issue, Meucci (2005) proposes the use of the 'expected return – expected shortfall' approach of Rockafellar and Uryasev (2002), in which the riskiness of an allocation is evaluated in terms of its ES (CVaR).

Subsequently, Meucci (2008a, 9–10; 2008b, 1–2) proposed a fully general BL approach called Entropy Pooling (EP) and showed that it could handle, e.g., non-normal reference risk models, non-linear views on various features (such as medians, volatilities, tail behaviors, tail co-dependence, copula, marginal distributions, joint distribution, et cetera), relative ranking and ordering information, simultaneous inputs from multiple users with different confidence levels, as well as scenario analysis, correlation stress-tests, external factors, and partial specifications. Here, the posterior incorporating all the inputs is defined as the one that minimizes the entropy relative to the prior and can be represented in terms of the same Monte Carlo scenarios as the reference model but with different probabilities.

The general theory of entropy pooling can be briefly explained as follows: First, let us consider a portfolio driven by an N-dimensional vector of risk factors $\boldsymbol{X}$. Second, let us denote the information currently available by $\mathcal{I}_t$, where t is the current time, and the time to the investment horizon is denoted by T. Thus, there exists a deterministic function P that maps the realizations of $\mathbf{X}$ and the information $\mathcal{I}_t$ into the price P_{t+T} of each security in the portfolio at the horizon, i.e., $P_{t+T} \equiv P(\mathbf{X}, \mathcal{I}_t)$. Regarding the reference model, it is assumed that there exists a model for the joint distribution of the risk factors, as represented by its probability density function. In other words, the risk model is denoted by $\mathbf{X} \sim f_{\mathbf{X}}$, which can be understood as the prior factor distribution in the original BL model. Investors use this model to optimize their positions by specifying a subjective index of satisfaction $\mathcal{S}$ (e.g., the mean-ES trade-off). Satisfaction depends both on the market distribution $f_{\mathbf{X}}$ through the prices and on the positions in the portfolio, represented by a vector $\mathbf{w}$. The optimal portfolio $\mathbf{w}^*$ can then be defined as

$$\mathbf{w}^* \equiv \underset{\mathbf{w} \in \mathcal{C}}{\arg\max} \{\mathcal{S}(\mathbf{w}; f_{\mathbf{X}})\}, \tag{2.3.4}$$

where $\mathcal{C}$ is a given set of investment constraints. (Meucci 2008b, 3.)

Views, in the most general case, are expressed on generic functions of the market $g_1(\mathbf{X}), \dots, g_K(\mathbf{X})$. These functions constitute a K-dimensional random variable whose joint distribution is implied by the reference model such that $\mathbf{V} \equiv \boldsymbol{g}(\mathbf{X}) \sim f_{\mathbf{V}}$. Since the views are statements on the variables that may clash with the reference model, the most detailed possible view specification in a stochastic environment is a complete, subjective joint distribution for those variables, that is, $\mathbf{V} \sim \tilde{f}_{\mathbf{V}} \neq f_{\mathbf{V}}$. In the original BL model, the views are statements on $\widetilde{\mathbb{E}}\{V_k\}$, i.e., the expectations of each of the V_k's according to the new distribution $\tilde{f}_{\mathbf{V}}$. However, in the entropy pooling approach, views are considered on a more general location measure $\widetilde{m}\{V_k\}$ that can be the expectation or the median. Thus, the views are set as

$$\widetilde{m}\{V_k\} \lesseqgtr m_k, \qquad k = 1, \dots, K, \tag{2.3.5}$$

and the values m_k can be determined exogenously. (Meucci 2008b, 4–5.)

Finally, the posterior distribution should satisfy the views without adding additional structure and be as close as possible to the reference model. Arguably, the relative entropy between a generic distribution $\tilde{f}_{\mathbf{X}}$ and a reference distribution $f_{\mathbf{X}}$ is a natural measure of the amount of structure in $\tilde{f}_{\mathbf{X}}$. Mathematically,

$$\mathcal{E}(\tilde{f}_{\mathbf{X}}, f_{\mathbf{X}}) \equiv \int \tilde{f}_{\mathbf{X}}(\mathbf{x})\left[\ln\tilde{f}_{\mathbf{X}}(\mathbf{x}) - \ln f_{\mathbf{X}}(\mathbf{x})\right]d\mathbf{x}. \tag{2.3.6}$$

Equation (2.3.6) also measures how distorted $\tilde{f}_{\mathbf{X}}$ is with respect to $f_{\mathbf{X}}$. If the two distributions coincide, the relative entropy is zero. By imposing constraints on $\tilde{f}_{\mathbf{X}}$, this distribution departs from $f_{\mathbf{X}}$ and relative entropy increases. Therefore, the posterior market distribution can be defined as

$$\tilde{f}_{\mathbf{X}} \equiv \underset{f \in \mathbb{V}}{\arg\min} \{\mathcal{E}(f, f_{\mathbf{X}})\}, \tag{2.3.7}$$

where $f \in \mathbb{V}$ stands for all the distributions consistent with the views statements. The posterior $\tilde{f}_{\mathbf{X}}$ follows by assuming that the investor has full confidence in her statements. If the confidence is less than full, the posterior distribution of the factors must shrink towards the reference factor distribution. In line with Meucci (2006), this can be achieved by opinion-pooling the reference model and the full-confidence posterior so that

$$\tilde{f}_{\mathbf{X}}^{c} \equiv (1-c)f_{\mathbf{X}} + c\tilde{f}_{\mathbf{X}}, \tag{2.3.8}$$

where the pooling parameter $c \in [0,1]$ represents the confidence level in the views. Hence, the full-confidence posterior is recovered if the confidence is total, whereas the reference risk model is recovered in the absence of confidence. (see Meucci 2008b, 4–7, 8–9, 19 for further technical details.)

2.4 Momentum as a financial-market anomaly

A large number of different patterns (or 'market anomalies') have been discovered in financial markets. These patterns can roughly be categorized as calendar anomalies (e.g., the January effect), fundamental anomalies (e.g., size and value effect), and technical anomalies (e.g., momentum and reversal). Many algorithms and techniques, such as machine learning, have been proposed to identify these patterns and generate profit by trading them (see Antonacci 2014; Gray & Vogel 2016; Li & Hoi 2016). Since we later assume that momentum could exist in the Russian sectoral indices possibly due to informational inefficiencies, let us now briefly focus on this particular anomaly.

Indeed, as noted in Moskowitz and Grinblatt (1999, 1249–1250), both investment theory and its application to investment management critically depend on understanding asset return persistence anomalies. Determining whether these anomalies are entrenched in behavior that more rational investors can exploit at low risk has profound implications for the view of market efficiency and optimal investment policy. The momentum effect, whose initial studies were published in the late 1980s and early 1990s (Israel & Moskowitz 2013, 286), can be defined as a continuation of the direction of prior asset returns (Griffin, Ji & Martin 2003, 2515; Wang & Kochard 2012, 53). As a further definition, cross-sectional momentum refers to the relative performance of a price series in relation to other price series, whereas time-series momentum means that past returns of a single price series are positively correlated with future returns. Momentum trading strategies generally utilize this tendency of assets with good (bad) recent performance to continue outperforming (underperforming) in the future by buying past winning assets and selling past losing assets. (Moskowitz, Ooi & Pedersen 2012, 228–229; Chan 2013, 133–134.)

Positive momentum effects have been found in securities (e.g., De Bondt & Thaler 1985; Conrad & Kaul 1988; Lo & MacKinlay 1988; Lehmann 1990; Jegadeesh & Titman 1993; Nijman, Swinkels & Verbeek 2004; Figelman 2007), international developed and emerging markets (Rouwenhorst 1998a; 1999; Griffin, Ji & Martin 2003; Erb & Harvey 2006; Asness, Moskowitz & Pedersen 2013), sectors and industries (Moskowitz & Grinblatt 1999) and asset classes (Blitz & van Vliet 2008; Moskowitz, Ooi & Pedersen 2012). Fama and French (2012) conclude that there is return momentum nearly everywhere and that spreads in average momentum returns decrease from smaller to bigger stocks (see also Hong, Lim & Stein 2000; Chen & Hong 2002, Grinblatt & Moskowitz 2004). Aït-Sahalia and Brandt (2001), in turn, find that the momentum effect, measured by lagged returns, has some predictive power.

Momentum is viewed as an anomaly because it is hard to explain within the standard asset pricing paradigm with a rational representative agent (Vayanos & Woolley 2013, 1). Furthermore, the finding of consistent time-series momentum (or 'trend' effect) in different instruments and markets seems to challenge the 'geometric random walk' hypothesis (see, e.g., Malkiel 2019), which in its most basic form implies that knowing whether a price went up or down in the past should not be informative about whether it will increase or decrease in the future. According to Moskowitz et al. (2012, 228–230), rejection of the random walk hypothesis does not necessarily imply a rejection of a more sophisticated notion of market efficiency with time-varying risk premiums, but the abnormal returns to time-series momentum also do not seem to be compensation for tail events or crash risk. Instead, the return to time-series momentum tends to be larger when the market experiences large up and down moves. Hence, time-series momentum could be a hedge for extreme events, making its large return premium even more perplexing from a risk perspective. On the other hand, the existence of a rational theory that can explain such findings may not be ruled out: For example, Docherty and Hurst (2018, 263) document a dynamic association between risk and the momentum premium by using return dispersion as a proxy indicator for structural shifts in the economy.

Similarly, Geczy and Samonov (2016, 32) argue that price momentum is dynamically exposed to market risk, conditional on the sign and duration of the trailing market state. Daniel and Moskowitz (2016), in turn, find that momentum strategies can experience infrequent and persistent strings of negative returns, in spite of their strong positive average returns across numerous asset classes. These 'momentum crashes' appear to be partly forecastable, and they occur in panic states, following market declines and when market volatility is high, and are contemporaneous with market rebounds. There also seems to be an asymmetry in the winner and loser exposure to market returns during extreme times. Although a reasonable explanation for equity momentum could be option-like payoffs of equities (see Merton 1974), the existence of the same phenomena and option-like features for momentum strategies in futures, bonds, currencies, and commodities makes such a theory less believable. (Daniel & Moskowitz 2016, 221–222, 242–243.)

Furthermore, Asness et al. (2013, 979–981) discover that funding liquidity risk is positively related to momentum globally across asset classes. In other words, momentum strategies do worse when liquidity declines, and their Sharpe ratios improve when liquidity improves (i.e., when borrowing is easier). Liquidity shocks also appear to increase correlations among momentum strategies significantly. Correspondingly, Pastor and Stambaugh (2003), as well as Sadka (2006), find that measures of liquidity are positively related to momentum in US individual stocks, but this link seems to be present also in other markets and asset classes. Therefore, liquidity risk may provide a partial but incomplete explanation for the positive risk premium associated with momentum. (see also Brunnermeier & Pedersen 2005; 2009; Pedersen 2009.)

A few papers also seek to examine whether momentum strategies remain profitable after considering transaction costs. For instance, Lesmond, Schill, and Zhou (2004) use theoretical models of transaction costs and find that standard relative strength strategies require frequent trading in disproportionately high-cost securities such that trading costs prevent profitable strategy execution (see also Korajczyk & Sadka 2004). Frazzini, Israel, and Moskowitz (2012), in turn, challenge these results by measuring the real-world transaction costs as well as price impact function facing an arbitrageur. The authors find that actual trading costs are less than a tenth as large as previous studies suggest. Similarly, Garleanu and Pedersen (2013) derive a closed-form optimal dynamic portfolio policy when trading is costly and find that strategies based on momentum predictors produce superior net returns relative to more naïve benchmarks. Furthermore, Israel and Moskowitz (2013, 275) find little evidence that momentum returns are significantly affected by changes in trading costs or institutional and hedge fund ownership over time. Hence, although there has generally been some degradation in profits to momentum strategies after their publication, arbitrageurs have not been able to eliminate all the profits (see, e.g., Schwert 2003; McLean 2010; McLean & Pontiff 2016). Liu and Longstaff (2004) argue that one reason for this may be due to margin requirements faced by risk-averse rational agents.

All in all, a vast amount of competing theories have been proposed to explain the short- and intermediate-term return continuation phenomenon (see, e.g., DeLong et al. 1990; Jegadeesh & Titman 1993; Chan, Jegadeesh & Lakonishok 1996; 1999; Barberis, Shleifer & Vishny 1998; Hong & Stein 1999; Hong et al. 2000; Holden & Subrahmanyam 2002; Scott, Stumpp & Xu 2003; Cooper, Gutierrez & Hameed 2004; Figelman 2007; Gutierrez & Prinsky 2007). Most of them are based on behavioral models, especially on the slow diffusion of news, which may cause initial underreaction (i.e., positive short-lag autocorrelations) and delayed overreaction (i.e., negative long-lag autocorrelations). For instance, Chan et al. (1996) argue that markets respond only gradually to new information, whereas Jegadeesh and Titman (1993, 66–67) show that the profits of momentum trading strategies cannot be attributed to a lead-lag effect resulting from delayed stock price reactions to information about a common factor. Instead, the evidence appears to be consistent with delayed price reactions to firm-specific information. (see also Lo & MacKinlay 1990; Jegadeesh & Titman 1995ab for possible explanations for the short-horizon return reversals and contrarian profits.)

Moskowitz and Grinblatt (1999, 1287), in turn, argue that if there are ‘hot’ and ‘cold’ initial public offering (IPO) markets, as well as ‘hot’ and ‘cold’ sectors of the economy, then investors may herd toward (away from) these hot (cold) industries and sectors, creating price pressure and return persistence. Alternatively, based on the theory proposed by Daniel, Hirshleifer, and Subrahmanyam (1997; 1998), investors may exhibit overconfidence and biased self-attribution in the presence of private and public information signals. This theory explains several of the stock return patterns that seem anomalous from the perspective of efficient markets with rational investors. It is also in line with the ones proposed by Hong and Stein (1999) and Barberis et al. (1998). In other words, they all argue that the momentum and reversal effects are part of the same phenomena, in that the momentum effect creates the subsequent reversal. When enough public information is disseminated, investors correct their initial errors, causing long-term return reversal. (see also Figelman 2007, 68–69.)

Although behavioral models are difficult to test empirically, there are several articles that indirectly support them. For example, Lee and Swaminathan (2000, 2065–2066) find that past trading volume, as measured by the turnover ratio, seems to predict both the persistence and magnitude of price momentum. In particular, high-volume stocks exhibit stronger momentum than low-volume stocks. The volume also ‘fuels’ momentum only for losers, and it helps information ‘diffusion’ only for winners. Consistent with the findings of Lee and Swaminathan (2000), Grinblatt and Han (2005) present a model of equilibrium prices in which a group of heterogenous investors is subject to the Prospect Theory and Mental Accounting (PT/MA) behavior (originally introduced by Thaler 1999). These investors have demand distortions that are inversely related to the unrealized profit they have experienced on a stock, and their demand functions distort equilibrium prices relative to those predicted by standard utility theory. Thus, a ‘winner’ stock that has been privy to prior good news has excess selling pressure relative to a ‘loser’ stock that has been privy to adverse information (cf. Shefrin & Statman 1985; Grinblatt & Keloharju 2001). If rational investors’ demand for a stock is not completely elastic, then such a demand disturbance caused by PT/MA tends to generate price underreaction to public information. This generates a spread between the stock’s fundamental value and market price. In equilibrium, past winners tend to be undervalued, and past losers tend to be overvalued. (Grinblatt & Han 2005, 313–315, 331)

In the model by Grinblatt and Han (2005), only the past returns’ pattern, combined with the past trading volume’s pattern, determines whether a stock has experienced an aggregate unrealized capital gain or loss. Stocks with paper capital gains will have higher average returns going forward than stocks with paper capital losses. Furthermore, stocks with large unrealized capital gains underreact to positive firm-specific news, while stocks with large unrealized capital losses underreact to negative firm-specific news. This is most likely to generate post-announcement drift (see Bernard & Thomas 1989; Frazzini 2006). Focusing on the ordinary common shares traded on the AMEX and NYSE exchanges, the authors find that there is a reversal of returns at both the short and long horizons but persistence in returns over the intermediate horizon. However, the momentum effect disappears when the capital gains overhang regressor (the difference between the current price and the aggregate cost basis) is included in the regression. (Grinblatt & Han 2005, 319–320, 333–336; see also Grundy & Martin 2001; George & Hwang 2004; Grinblatt & Moskowitz 2004 for related studies.)

In contrast, Vayanos and Woolley (2013, 1–3) propose an institutional theory of momentum and reversal based on flows between investment funds. According to this delegated portfolio management model, flows are triggered by changes in fund managers' efficiency, which investors either observe directly or infer from past performance. Momentum arises because rational prices underreact to expected future flows or if flows exhibit inertia due to slow-moving capital. Reversal, in turn, arises because flows push prices away from fundamental values. Besides momentum and reversal, flows generate co-movement, lead-lag effects, and amplification, with these being larger for high-idiosyncratic-risk assets. It is further assumed that rational investors buy assets whose expected returns have decreased because of the so-called 'bird-in-the-hand' effect. In other words, although the price is expected to drop even further in the short run, the investor buys an underpriced asset to hedge against a reduction in the mispricing (cf. the Intertemporal Capital Asset Pricing Model by Merton 1973).

Vayanos and Woolley (2013, 28–29, 31–33) find that momentum strategies, which are imperfect approximations of the optimal strategy, can profitably exploit aspects of flow-generated mispricing. Empirical support for their theory can also be found, e.g., from Barberis and Shleifer (2003), Coval and Stafford (2007), Dasgupta, Prat, and Verardo (2011), as well as from Lou (2012). For example, Coval and Stafford (2007) find that mutual funds experiencing large redemptions are likely to reduce or eliminate their existing stock positions. As these asset 'fire sales' depress the stock prices, they also suppress the performance of other funds holding those stocks. In reverse, the same situation occurs for stocks disproportionately held by well-performing mutual funds with large capital inflows. As the stocks suffer suppression or elevation of their prices through no fault or merit on their own, the prices often mean-revert after the mutual fund selling or buying pressure is over.

In addition to interday momentum strategies, research has also been done on strategies based on intraday momentum trading. For example, it is well-acknowledged that index composition changes induce momentum in stocks that are added to or deleted from the index (see, e.g., Shankar & Miller 2006; Cheng & Madhavan 2009; Shum et al. 2015).

Although such momentum could potentially last several days, the tests conducted by Chan (2013, 162–163) suggest that the drift horizon has been reduced to intraday, at least in the US stock markets. The actions of large funds can also trigger intraday momentum: the daily rebalancing of leveraged ETFs near the market close causes short-term momentum in the underlying index in the same direction as the market return from the previous close. Furthermore, at the shortest possible time scale, the imbalance of the bid and ask sizes, the changes in order flow or non-uniform distribution of stop orders can all induce momentum in prices (see, e.g., Osler 2000; 2003; Lyons 2001; Maslov & Mills 2001; Hellström & Simonsen 2006). Alternatively, high-frequency traders may cause artificial return persistence by market manipulation. (see, e.g., the books by Harris 2002; Durbin 2010; Sinclair 2010; Arnuk & Saluzzi 2012 for more discussion on practical HFT strategies and the pitfalls of day trading tactics.)

Behavioral models may be utilized again to explain why, e.g., order flow information can act as a predictor of price movements. More precisely, it can be assumed that market makers set the bid-ask prices based on the order flow information. Then, if a major hedge fund learns about a piece of breaking news, their algorithms will submit large market orders of the same sign in a split second. A market maker monitoring the order flow will deduce that such large one-directional demands indicate the presence of informed traders, and they will immediately adjust their bid-ask prices to protect themselves. (see Chan 2013, 156, 165–168.)

In conclusion, it is worth noting that, especially in the case of the shorter-term news-driven momentum, the duration over which momentum remains in force gets progressively shorter as more traders catch on to it. This is understandable if price momentum is viewed as generated by the slow diffusion of information. As more investors learn about the information faster and earlier, the diffusion–and thus, momentum– also ends sooner, which, in turn, constantly shortens the holding period. Thus, the performance of the momentum strategy in question may eventually start to decay without a visible explanation. On the other hand, since there are two common types of exit strategies for momentum strategies (namely time-based and stop-loss), risk management is often relatively easy, and the loss of a momentum position is always limited. Finally, momentum models tend to thrive on so-called ‘black swan’ events (Taleb 2010), meaning that the thicker the tails of the returns distribution curve, or the higher its kurtosis, the better that market typically is for momentum strategies. (see Chan 2013, 61, 151, 153–154.)

2.5 Risk management and risk measures

In this subchapter, we briefly discuss risk measures and portfolio risk management. As it is known, the main objective of portfolio selection is the construction of portfolios maximizing expected returns at a certain level of risk. Correspondingly, it is well established that asset returns are not normally distributed, and thus the mean and variance alone do not fully describe the characteristics of the joint asset return distribution. The variance of the portfolio, for example, cannot solely capture and quantify several risks and undesirable scenarios. Moreover, using variance as the only risk measure in the portfolio optimization means that outcomes above the expected portfolio return are considered as risky as outcomes below the expected portfolio return. This is counterintuitive because investors are more likely to be concerned about outcomes that fall short of expectations rather than outcomes that exceed them. Finally, the assumption that asset returns follow elliptical probability distributions is hazardous during financial crises. As a result, the tail risk has become a critical statistical determinant of risk management policies. (Fabozzi et al. 2010, 342; Pachamanova & Fabozzi 2016, 291–292.)

Risk measures can be broadly categorized as either dispersion or downside measures. Dispersion measures, such as mean standard deviation, mean absolute deviation, and mean absolute moment, are measures of uncertainty, and they treat both positive and negative deviations from the mean as equally risky. Thus, overperformance relative to the mean is penalized as much as underperformance. In contrast, the objective of downside risk measures in a portfolio context is to maximize the probability that the portfolio return is above a certain acceptable minimum level (aka the benchmark or disaster level). Despite their theoretical appeal, the downside (or safety-first) risk measures can be computationally burdensome. It can also be challenging to aggregate downside risk measures of individual securities into portfolio downside risk measures because their calculation requires knowledge of the entire joint distribution of security returns. Hence, the estimation risk of downside measures tends to be greater than for standard mean-variance approaches. (Fabozzi et al. 2010, 343–357; see also Ortobelli et al. 2005; Teplova & Shutova 2011, 160 for further discussions.)

Due to space constraints, we focus on two well-known downside risk measures: Value-at-Risk (VaR) and Conditional Value-at-Risk (CVaR). However, it is worth emphasizing that the literature related to risk measures is vast. For example, draw-down optimal portfolios, which are closely related to CVaR-optimal portfolios, are discussed in Chekhlov, Uryasev, and Zabarankin (2000; 2005), as well as in Krokhmal, Uryasev, and Zrazhevsky (2002). Litterman (1996) and Szegö (2004) provide comprehensive illustrations of how various risk measures can be applied in asset management. Gourieroux, Laurent, and Scaillet (2000) and Gourieroux and Liu (2006), in turn, study the sensitivity of risk measures with respect to portfolio allocation, as well as to parameters representing risk aversion or pessimism. Furthermore, Kusuoka (2001) and Acerbi (2002; 2007) introduce a general class of distribution-based (or law invariant) risk measures called 'spectral risk measures.' Finally, dynamic (i.e., time-varying) risk measures are proposed, among others, by Cvitanic and Karatzas (1998), Weber (2003), Frittelli and Rosazza Gianin (2004), and Riedel (2004).

Let us start with Value-at-Risk, which is often utilized in risk management as well as in portfolio allocation. VaR is related to the percentiles of loss distributions, and it measures the predicted maximum portfolio loss at a specified probability level over a specific time horizon. Hence, a portfolio allocation aiming to reduce losses usually reduces the VaR. Alternatively, VaR can be used to both track and report the market risk exposure of trading portfolios. Mathematically, Value-at-Risk can be defined as follows:

$$\mathrm{VaR}_{\propto}(\mathrm{R_p}) = \min\left\{\mathrm{R}\middle|\mathrm{P}(-\mathrm{R_p} \leq \mathrm{R}) \geq \alpha\right\}, \quad (2.5.1)$$

where P denotes the probability function, and $\mathrm{R_p}$ denotes the portfolio returns. Typical values for the confidence level α are 90%, 95%, and 99%. For example, the 95% VaR for a portfolio is the value R such that the probability that the possible portfolio loss exceeds R is less than the risk level $(1-\alpha) = 5\%$. (Fabozzi et al. 2010, 348–350; see also Pachamanova & Fabozzi 2016, 191–193, 292–293 for more details on the portfolio VaR minimization problem.) Different methods have been proposed to calculate the Value-at-Risk of a portfolio containing either stocks (e.g., Hopper 1996; Johansson, Seiler & Tjarnberg 1999), bonds (e.g., Golub & Tilman 1997; Phelan 1997; Hull 2018), or derivatives (e.g., Jorion 1997; Singh 1997; Kupiec 1998; Wilmott 1998; Ahn et al. 1999; El-Jahel, Perraudin & Sellin 1999; Linsmeier & Pearson 2000). These include inter alia, the variance-covariance, and delta approaches, as well as historical and Monte Carlo simulation methods (see Lyuu 2004, 276–278, 361–374, 474–478 for a general overview). The computation of VaR when returns are not normally distributed has also been studied, e.g., by Duffie and Pan (1997), Venkataraman (1997), Hull and White (1998), and Neftci (2000). Finally, Beder (1995) shows that the VaR is highly dependent on parameters, data, assumptions, and methodology.

There are a couple of reasons why VaR is considered an inconvenient risk measure. Firstly, a risk measure should be *coherent*, which, by definition, means that it should satisfy the four axioms of translation invariance, subadditivity, positive homogeneity, and monotonicity (Artzner et al. 1999, 210; see also Acerbi & Tasche 2002; Krokhmal 2007; Shapiro 2013). Value-at-Risk, however, is not subadditive. This means that the risk as measured by the VaR of a portfolio of two funds can be higher than the sum of the risks of the two individual portfolios, which is in contrast to the diversification effect stating that a more diversified portfolio should have a lower risk (Fabozzi et al. 2010, 348–350; see also Frittelli & Rosazza Gianin 2002; Föllmer & Schied 2002; Song & Yan 2009). On the other hand, the usefulness of diversification has already been questioned in situations where the asset returns are dependent (see, e.g., MacMinn 1984; Hong & Herk 1996;

Kijima 1997). Furthermore, Ibragimov (2009, 565–566, 572), who focuses on portfolio diversification and value-at-risk analysis under heavy-tailedness, shows that the stylized fact that portfolio diversification is preferable is reversed for extremely heavy-tailed risks or returns. Thus, requiring subadditivity in such a case would be unnatural (see also Ibragimov & Walden 2007, 2567).

Further arguments against the over-emphasis of subadditivity have also been provided. For example, Danielsson et al. (2005) explore the potential for violations of VaR subadditivity theoretically and by simulations and find that VaR is subadditive when the risks do not have extremely heavy tails (see also Kou, Peng & Heyde 2013, 406). Heyde, Kou, and Peng (2006), in turn, argue against subadditivity from an axiomatic point of view and propose a new, data-based risk measure called 'natural risk statistic' that is characterized by a weaker set of axioms. Garcia, Renault, and Tsafack (2007) claim that when the tail thickness is responsible for violation of subadditivity, eliciting proper conditioning information can restore VaR rationale for decentralized risk management. Finally, Dhaene et al. (2008) support the use of VaR from the perspective of solvency capital requirements and show that the subadditivity condition can lead to the undesirable situation where the shortfall risk increases by a merger.

Another criticism of the Value-at-Risk is that it does not provide a measure of the magnitude of losses suffered above the threshold (Diaz et al. 2019, 3). To illustrate, if there are 95 sunny days and five rainy days, we are usually not interested in the least amount of rain to expect on the five rainy days but instead would like to estimate the average rainfall in those days (Pfaff 2016, 39). Analogously, it is unlikely that an investor will be indifferent between two portfolios with identical expected returns and VaR when the return distribution of one portfolio has a short left-tail, and the other has a long left-tail (Fabozzi et al. 2010, 348–350).

The Conditional Value-at-Risk (CVaR) risk measure directly addresses the two above-mentioned issues, and it can be represented by using the following formula:

$$\mathrm{CVaR}_{\alpha}(\mathrm{R_p}) = \mathrm{E}\left(-\mathrm{R_p} \mid -\mathrm{R_p} \geq \mathrm{VaR}_{\alpha}(\mathrm{R_p})\right). \tag{2.5.2}$$

Therefore, CVaR measures the expected amount of losses in the tail of the distribution of possible portfolio losses exceeding the portfolio VaR for a given confidence level. For continuous loss distributions, CVaR is identical to the Expected Shortfall (ES) (see Acerbi & Tasche 2002, 1496). Since the axiom of subadditivity is fulfilled, both ES and CVaR are coherent risk measures implying that the CVaR (or ES) of a portfolio cannot be greater than the sum of the individual risks. Moreover, when the CVaR is employed for portfolio optimization, less concentrated portfolios are typically achieved as a result than in the case of VaR. From an optimization point of view, it can also be shown that CVaR is a convex function, whereas VaR may be non-convex. Naturally, the non-convexity can yield portfolio solutions pertinent to local (and not global) optima. On the other hand, an outright optimization regarding CVaR/ES is complicated from a numerical point of view if one uses such a formulation where the CVaR depends on the VaR. (Carl, Peterson & Boudt 2010, 18; Fabozzi et al. 2010, 350–353; Pfaff 2016, 39, 234.)

Instead of utilizing the CVaR function whose definition involves the VaR function explicitly, it is possible to consider a more straightforward auxiliary objective function with better computational properties. Rockafellar and Uryasev (2002) were the first to propose this kind of approach. Pachamanova and Fabozzi (2016, 294–297) show further that when the joint probability density function of the returns for the stocks in the portfolio is represented in a set of scenarios, the minimization of CVaR becomes a very tractable optimization problem, which, due to its linear structure, can be implemented with most solvers (see, e.g., Tütüncü, Toh & Todd 2003). Alternatively, as demonstrated by Jurczenko and Teiletche (2015, 148–149), the CVaR of the portfolio returns can be written as follows:

$$\mathrm{CVaR}_{\varepsilon}(\mathbf{w}) = -\mu_{\mathrm{p}} - \sigma_{\mathrm{p}}\mathrm{E}[\mathrm{Z}|\mathrm{Z} \leq \mathrm{G}^{-1}(\varepsilon)], \tag{2.5.3a}$$

with

$$\mathrm{E}[\mathrm{Z}|\mathrm{Z} \leq \mathrm{G}^{-1}(\varepsilon)] = \frac{1}{\varepsilon}\int_{0}^{\varepsilon} \mathrm{G}^{-1}\,(\mathrm{s})\,\mathrm{ds}, \tag{2.5.3b}$$

where μ_{p} and σ_{p} are the portfolio's mean return and volatility, respectively, ε is the risk level, $\mathbf{w} = (w_1, \dots, w_n)$ is a portfolio made up of n assets, $\mathrm{Z} = (\mathrm{R}_{\mathrm{p}} - \mu_{\mathrm{p}})/\sigma_{\mathrm{p}}$ is a zero-mean unit-variance random variable with distribution function $\mathrm{G}(\cdot)$, and $\mathrm{G}^{-1}(\cdot)$ is the quantile function associated with the portfolio's return distribution. By using this formulation, it can thus be seen that the higher expected return μ_{p} reduces CVaR, whereas higher volatility σ_{p} increases CVaR.

With respect to efficient frontiers, De Giorgi (2002, 2, 12, 20–21) shows that under the assumption of multivariate Gaussian distributed returns, the set of efficient portfolios under VaR and CVaR is a subset of the set of efficient portfolios under the standard deviation. Therefore, (μ, σ)-portfolio selection could be inefficient under VaR or CVaR, but the opposite never occurs. Fabozzi et al. (2010, 354–356) continue by noting that when loss distributions are non-normal, the mean-variance and the mean-CVaR approaches may differ significantly (see also Dowd 1999, 65–66; Krokhmal, Palmquist & Uryasev 2001, 19; Würtz et al. 2015a, 243–259 for more discussions on risk-return analyses under different risk measures). Finally, Alexander and Baptista (2004) analyze the portfolio selection implications of imposing either a VaR or a CVaR constraint on the mean-variance model. The authors show that for a given confidence level, a CVaR constraint is tighter than a VaR constraint if the CVaR and VaR bounds coincide. As a result, a CVaR constraint is more effective as a tool to control slightly risk-averse agents, but in the absence of a risk-free asset, it has a twisted effect in that it is more likely to force highly risk-averse agents to select portfolios with larger standard deviations. However, when the CVaR bound is appropriately larger than the VaR bound, or when a risk-free asset is present, a CVaR constraint 'dominates' a VaR constraint as a risk management tool.

Overall, since the advent of the 2008 financial crisis, risk-based investment strategies and various risk budgeting policies have attracted increasing attention (see, e.g., Roncalli 2013; 2014 and references therein). However, in most applications, it is common to determine the capital allocations so that all portfolio components contribute to the same extent to its volatility. A novel exception is provided by Jurczenko and Teiletche (2015, 147–148, 150–153), who propose a generalized risk-based asset allocation framework, where volatility, correlation, asymmetry, valuation, tail, and illiquidity risks can all be incorporated. In their approach, carry is used as a measure of expected returns (see, e.g., Ilmanen 2011; Koijen et al. 2018), and Expected Shortfall is adjusted for the non-Gaussian features of the distribution of returns by utilizing the Cornish-Fisher approximations (see, e.g., Johnson, Kotz & Balakrishnan 1994). The empirical estimates of covariance, co-skewness, and co-kurtosis are then corrected with the help of moving average models, and a new expression for the individual risk contributions to the total ES is derived by isolating illiquidity risk. (see also Boudt, Peterson & Croux 2008 for an alternative way to include the higher-order moments in the computation of the ES.)

When considering the practical usefulness of the risk measures, there are two additional aspects worth mentioning, namely the international regulatory framework and robust risk measurement procedures. Firstly, although the Basel Committee on Banking Supervision (2019) has changed the measure to determine market risk capital from 99 percent VaR to 97.5 percent ES, it is not clear which method should be employed to evaluate the goodness of the proposed risk measure. Indeed, designing a backtesting method for the expected shortfall is not as straightforward as VaR's case since ES does not satisfy the elicitability property (see Weber 2006; Gneiting 2011). Hence, an appropriate scoring function that this risk measure potentially minimizes does not exist. On the other hand, Fissler and Ziegel (2016) prove that ES is elicitable if combined with VaR. (see also Diaz, García-Donato & Mora-Valencia 2019, 2–4.)

As for robustness, Mausser and Rosen (2000) show that ES is less sensitive than VaR to sampling noise in a simulation-based environment and, therefore, tends to provide more robust risk analytics. In contrast, Yamai and Yoshiba (2002) find that ES needs a larger sample size than VaR for the same level of accuracy. Finally, Cont, Deguest, and Scandolo (2010) argue that most parametric estimation procedures for VaR and ES lead to non-robust estimators. Moreover, contrary to the historical VaR, the sensitivity of historical ES is unbounded, which explains why this risk measure shows high sensitivity to 'outliers.' Hence, there seems to be a conflict between the subadditivity of a risk measure and the robustness of its estimation procedure (see also Kou et al. 2013, 402, 406).

Let us conclude by demonstrating how downside risk measures can be combined with volatility and dependence modeling. Concerning the former, it is well-known that several stylized facts characterize asset returns (see, e.g., Bollerslev et al. 1992; 1994; Campbell & Hentschel 1992). Correspondingly, a large number of conditional variance models seeking to capture the characteristics of financial data have been proposed since the seminal papers of Engle (1982) and Bollerslev (1986) (see, e.g., Hansen & Lunde 2005, 873; Poon & Granger 2003 for extensive reviews). Indeed, the distinctive feature of models like the Generalized ARCH is that they recognize that volatility is not constant (i.e., it clusters, possesses long memory, and mean-reverts) (Hull 2012, 205–206; Sinclair 2013, 36–40; see also Cont 2001; Poon 2005 for detailed expositions). Various GARCH-based models have also been quite successful in predicting conditional means of asset returns (see, e.g., Beach & Orlov 2007, 150–151), and the variance and covariance forecasts based on a multivariate GARCH process have been shown to be essential in portfolio construction (see Pojarliev & Polasek 2003, 114). On the other hand, Lehar, Scheicher, and Schittenkopf (2002) show that, although the GARCH model dominates the Hull and White (1987) stochastic volatility in option pricing, there are no significant differences between the models when they are used for risk management (i.e., for calculating VaR forecasts) (see also Musiela & Rutkowski 2005, 273). Finally, a significant disadvantage of volatility measures derived from different parametric models is that they may not be robust to misspecification (Anatolyev 2008, 58).

When considering the modifications and extensions of the standard GARCH model, it is worth noting that simple models are often preferred in the forecasting literature to more complicated models with a higher degree of parametrization (see Low et al. 2018, 264–265 and references therein). The reason is that complex econometric models can suffer from issues such as data mining, poor performance out-of-sample, and failure to produce meaningful profitability in a portfolio management context (Kritzman, Page & Turkington 2010, 33). In practical applications, it is also frequently observed that models with smaller order sufficiently describe the data, and GARCH(1,1) is adequate in most cases (Hull 2012, 218). However, the inability of the basic GARCH model to serve as a true data generating process and as an accurate forecasting tool has also been disclosed. For example, Franke et al. (2015, 293–295) find that the GARCH(1,1) model greatly overestimates the variance for the long series of S&P 500 returns and provides poor longer horizon forecasts (see also Starica 2003; Mikosch & Starica 2004).

A particularly useful volatility model containing several variations of the discrete-time univariate GARCH(p, q) model is the Asymmetric Power ARCH (APARCH), which was initially proposed by Ding, Granger, and Engle (1993). Following the same notation as in Pfaff (2016, 117–120), the model's conditional variance equation, h_t, can be defined as follows:

$$\epsilon_t = \eta_t h_t, \tag{2.5.4a}$$

$$\eta_t \sim \mathcal{D}_v(0,1), \tag{2.5.4b}$$

$$h_t^\delta = \alpha_0 + \sum_{i=1}^{q} \alpha_i(|\epsilon_{t-i}| - \gamma_i \epsilon_{t-i})^\delta + \sum_{j=1}^{p} \beta_j h_{t-j}^\delta\,, \tag{2.5.4c}$$

with parameter restrictions $\alpha_0 > 0$, $\delta \geq 0$, $\alpha_i \geq 0$, $i = 1, \dots, q$, $-1 < \gamma_i < 1$, $i = 1, \dots, q$, and $\beta_i \geq 0$, $i = 1, \dots, p$. Here, ϵ_t is the error process, η_t denotes a random variable with distribution $\mathcal{D}$ with expected value zero and unit variance. Additional parameters of this distribution (e.g., the skewed Student's t distribution) are subsumed in the parameter vector $\mathscr{v}$. The exponent parameter δ represents a Box-Cox transformation of h_t, and since it is generally defined for non-negative values, the long memory characteristic (i.e., slow decaying autocorrelations) often encountered for absolute or squared daily return series can be taken explicitly into account. Moreover, potential asymmetries with respect to positive and negative shocks are reflected in the coefficients $\gamma_i, i = 1, \dots, q$. Hence, the APARCH model encompasses the following special cases: the plain GARCH, if $\delta = 2$ and $\gamma_i = 0$ (see Bollerslev 1986); GJR-GARCH, if $\delta = 2$ (see Glosten, Jagannathan & Runkle 1993); TS-GARCH, if $\delta = 1$ and $\gamma_i = 0$ (see Taylor 1986; Schwert 1990); Threshold-ARCH, if $\delta = 1$ (see Zakoian 1994); N-ARCH, if $\gamma_i = 0$ and $\beta_j = 0$ (see Higgins & Bera 1992); as well as Log-ARCH, if $\delta \to 0$ (see Geweke 1986; Pentula 1986). Naturally, the desired risk measures can be deduced from the recursively computed volatility forecasts together with the quantiles of the assumed distribution for the error process (Pfaff 2016, 120).

Another well-known extension not included in the above specification is the class of Exponential GARCH (EGARCH) models proposed by Nelson (1991). Here, the 'leverage effect' often observed in empirical studies (see, e.g., Black 1976; Figlewski & Wang 2000; Bollerslev, Litvinova & Tauchen 2006; Inkaya & Yolcu Okur 2014; Aboura 2015; Sun & Wu 2018) is captured as follows:

$$\log(h_t) = \alpha_0 + \sum_{i=1}^{q} \alpha_i g(\eta_{t-i}) + \sum_{j=1}^{p} \beta_j \log\left(h_{t-j}\right), \tag{2.5.5}$$

where the function $g(\eta_t)$ is defined as $g(\eta_t) = \theta\eta_t + \gamma[|\eta_t| - \mathrm{E}(|\eta_t|)]$. Thus, the impact of the error variables takes a value of $\alpha_i(\theta + \gamma)$ for a positive shock and $\alpha_i(\theta - \gamma)$ for a negative one, whereas the coefficient γ captures the ARCH effects. Since this approach uses the logarithmic form, there is no need for non-negativity constraints on the parameter space, α_i, $i = 1, \dots, q$ and β_j, $j = 1, \dots, p$. (Pfaff 2016, 119.) However, the existence of unconditional moments depends on the choice of the distribution of the innovations, which is an undesirable property of the EGARCH models (Franke, Härdle & Hafner 2015, 285–287). It has also been found that EGARCH often over-weights the effects of larger shocks on volatility and thus results in poorer fits than standard GARCH models (see, e.g., Engle & Ng 1993).

The applicability of the above volatility models in the risk management context can be illustrated by conducting a backtest of the conditional Expected Shortfall for a financial time series. This is done in Figure 2.2 using the daily returns of Lukoil Oil Company. The total return index data denoted in Russian rubles is downloaded from the Thomson Reuters Datastream, and it covers the period from May 9, 1997, to May 11, 2017 (5220 daily observations in total). The daily compound losses are computed as percentages. Here, a $\mathrm{GARCH}(1,1)$ model with a Student's t-distributed innovation process is estimated first. Then, the one-step-ahead forecast of the conditional standard deviation is computed as well as the ES for the confidence level of 99 percent. In line with Pfaff (2016, 128–130), a moving window with 1000 observations is utilized, and the mean equation of the GARCH model is omitted.

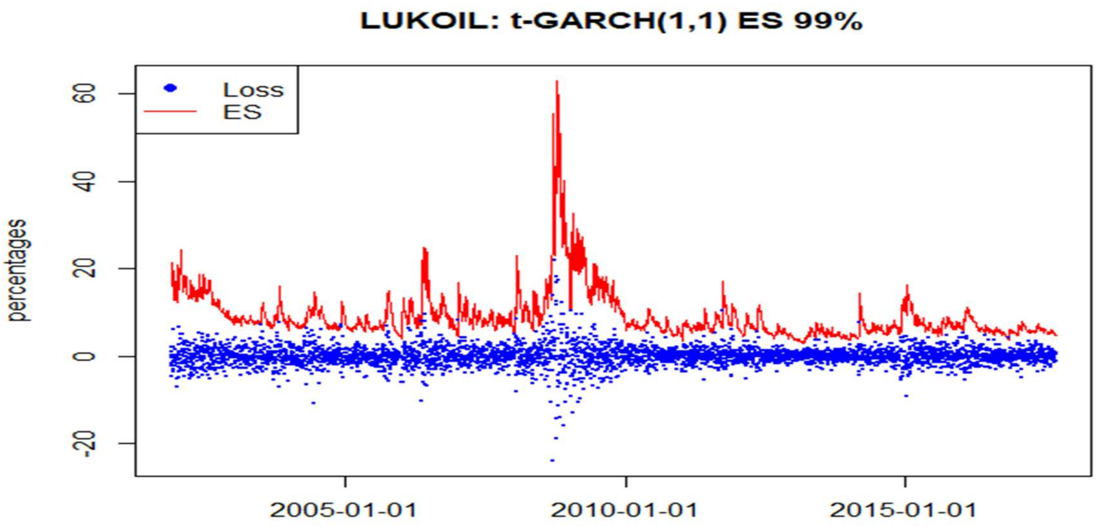

Figure 2.2 Graphical comparison of Lukoil's actual daily losses and the conditional ES for a 99 percent confidence level

When comparing the conditional risk measures with the actual losses, it can be seen that this risk measure captures the spikes of the daily losses rather well: Its level, which peaked at over 60 percent between mid-2008 and early 2009, decreases rapidly during the more quiet periods. Moreover, during our backtest simulation, only about five violations occurred. Such a result can be considered conservative, given that roughly 15 exceedances can be expected (Pfaff 2016, 130; see also Ruppert & Matteson 2015, 557–559, 561–565 for more discussion on estimating VaR and ES using ARMA+GARCH models for a portfolio of assets.)

GARCH models are also widely applicable when quantifying extreme market risk and extreme financial scenarios. In such a case, the Extreme Value Theory (EVT) provides a solid statistical modeling framework of interest (see Singh & Allen 2017, 199–201). For example, McNeil and Frey (2000) and Singh, Allen, and Powell (2013) propose a dynamic forecasting method that uses a GARCH-type process along with the Peaks Over Threshold (POT) to model VaR and ES reacting to current market fluctuations. This approach is depicted in Figure 2.3: It visualizes the one percent Value-at-Risk and 2.5 percent Expected Shortfall forecasts for the same daily percentage log-returns of Lukoil as in the previous case. Here, an AR(1)-APARCH(1,1) model is first used to predict the one-day-ahead mean expected return and volatility. The residuals of the first step are then fitted to the Generalized Pareto Distribution (GPD) to get quantile values for the final risk measure calculations. As in Singh and Allen (2017, 221–224), a 90 percent quantile level is chosen as the threshold, and a moving window is set to 1000 observations. (see also Würtz, Setz & Chalabi 2017b, 34 for more technical details about the algorithm used.)

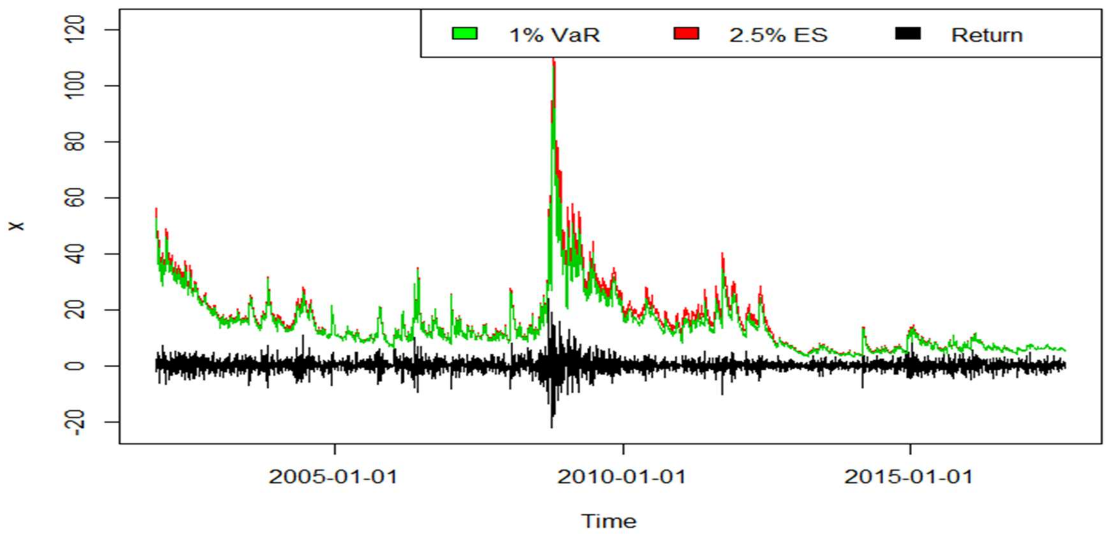

Figure 2.3 Dynamic-EVT 99%-VaR and 97.5%-ES forecasts against the actual daily percentage returns of Lukoil

As can be seen, both the ES at the 97.5 percent confidence level and equally the VaR at the 99 percent confidence level (see Acerbi & Szekely 2014, 3) greatly overestimate the risks in the aftermath of the 1998 Russian financial crisis as well as during the global financial crisis in 2008. Another evidence of conservative risk estimates is that the plot shows no clear VaR or ES exceedances (see, e.g., Emmer, Kratz & Tasche 2015; Kratz, Lok & McNeil 2018 for more discussion on backtesting methodologies for correct exceedances for VaR and ES forecasts). On the other hand, the figure does show how the risk measures follow the actual returns and adjust themselves to changing volatility in a time series. Similarly, Diaz et al. (2019, 14–15, 18–19) find that the GPD model tends to overpredict risk when VaR is calculated at confidence levels higher than 97.5 percent.

Finally, let us turn to copulas, which can be utilized to model a correlation structure between two or more variables, regardless of the shapes of their probability distributions. As is well known, the (Pearson) coefficient of correlation measures *linear* dependence between two variables, but there are many other ways in which two variables can be related. By definition, copulas are functions that join or 'couple' multivariate distributions to their one-dimensional marginal distribution functions. Equivalently, copulas are multivariate probability distribution functions whose marginals are uniform on the interval $(0,1)$. Since the specification of marginal distributions can be separated from the dependence structure, copulas provide a way of studying scale-free measures of dependence. (see, e.g., McNeil, Frey & Embrechts 2005, 184–191; Chen & Fan 2006, 126; Balakrishnan & Lai 2009, 33; Wilson & Ghahramani 2010, 7–8 for more information on the mathematics behind copula functions.)

Without going into technical details, it is worth mentioning that, as in the case of volatility models, numerous copula-based models have been introduced in the literature. For example, the notorious Gaussian copula (see, e.g., MacKenzie & Spears 2014) is asymptotically independent in both tails, meaning that if one goes far enough into the tail, extreme events appear to occur independently in each margin regardless of how high a correlation is chosen. This problem can be overcome by using the Student's t copula, which has both upper and lower tail dependence of the same magnitude. Hence, it can provide a better description of the joint behavior of market variables than the Gaussian copula because the correlations between market variables tend to increase in extreme market conditions. (see McNeil et al. 2005, 211–212, 218; Hull 2012, 244–246.) As a further classification, both the Gaussian and Student's t copulas are part of the elliptical copula family, whereas, e.g., the Gumbel, Clayton, Frank, and Joe copulas belong to the large family of Archimedean copulas (see, e.g., Nelsen 2006; Cherubini et al. 2012, 30–35; Cherubini, Gobbi & Mulinacci 2016, 23–25; Hofert et al. 2018 for more discussion on different copula families). For instance, as opposed to the Gumbel copula, the Clayton copula is particularly helpful in capturing *lower* tail dependence and thus suitable for financial risk modeling (see Singh & Allen 2017, 235–241).

Copula models have also been flexibly combined with volatility modeling. Indeed, the first GARCH-copula approach can be traced back to Rockinger and Jondeau (2001), and the combination of models they proposed has since been used not only for the modeling of portfolio risk but also in the domain of financial derivatives. For example, Embrechts and Dias (2003), Micocci and Masala (2005), Jondeau and Rockinger (2006b), Patton (2006ab), Chiou and Tsay (2008), and Hsu et al. (2008) have utilized this kind of approach in their empirical research. Furthermore, Wang et al. (2010) apply a GARCH-EVT-copula model to evaluate the risk of foreign exchange portfolio and find that the optimal investment allocations are similar across different copulas and confidence levels, but t-copula and Clayton copula better portray the correlation structure of multiple assets than normal copula. Wang, Jin, and Zhou (2010) continue by pointing out that the Student-t provides a better estimation of the VaR than the Gaussian and Clayton copulas.

Huang and Hsu (2015), in turn, find that GARCH-EVT-copula models can generate significant economic gains when employed in constructing optimal out-of-sample Minimum CVaR portfolios, especially if the rebalancing interval is short.

Based on these earlier studies, Sahamkhadam, Stephan, and Östermark (2018) find that a Certainty Equivalence Tangency (CET) portfolio, which maximizes the Sharpe ratio and is based on ARMA-GARCH-EVT-copula forecasts, outperforms the benchmark portfolio based on historical returns in terms of wealth accumulation. Interestingly, the authors also conclude that elliptical copulas reduce the return volatility of the Min-CVaR portfolio better than Archimedean copulas. However, while the use of such forecasting models can reduce the portfolio risk measured either by the portfolio return volatility or the first percentile of portfolio returns, it also reduces the likelihood of high portfolio returns. Therefore, in terms of risk-modeling strategies, the benefits may come at the cost of reducing the 'opportunities' in the upper tail of the asset returns as well, thus implying lower performance. (see Sahamkhadam et al. 2018, 503–505.) In contrast, Chu and Satchell (2018, 249–250) show that due to misspecifications, the normal copula may sometimes outperform the Clayton or Gumbel copula even in terms of economic gains if one considers a specific 'style investing' problem where an investor with the Constant Relative Risk Aversion (CRRA) utility allocates wealth between the 'growth' and 'value' indices.

Recently, the so-called *vine copulas* have received widespread attention. This is because they have been shown to increase the flexibility of multivariate copula models beyond the scope of elliptical and Archimedean copulas, which may be inadequate models to describe the dependence characteristics of real data applications (see, e.g., Fischer et al. 2009). In brief, vine copulas use (conditional) bivariate copulas as 'pair copula' building blocks to describe a multivariate distribution. A set of linked trees (i.e., the 'vine') depicts a vine copula's factorization of the multivariate copula density function into the density functions of its pair copulas. As a pair copula construction, vine copulas allow different structural behaviors of pairs of variables to be modeled properly. (Czado, Brechmann & Gruber 2013, 17–21.) If the trees embedded one in the other have at least one 'node' in common, the vine is called 'Regular' (R-vine). Regular vines, which can be depicted in a graph-theoretic model, can be considered as specializations of Cantor trees for which all constraints are two-dimensional or conditional two-dimensional. Further restrictions make it possible to distinguish two different categories of regular vines, namely the Canonical vines (C-vines) and Drawable vines (D-vines). (Cherubini et al. 2012, 44; Singh & Allen 2017, 242; see also Bedford & Cooke 2001; 2002; Aas et al. 2009; Czado 2019 for more mathematical details.) C-vine copulas may be used if there is a set of pivotal variables such as stock indices (see, e.g., Heinen & Valdesogo 2009), whereas D-vine copulas are particularly attractive to model variables with temporal order (see, e.g., Brechmann & Czado 2015).

Several studies have applied vine copulas to financial risk management. For example, Brechmann and Czado (2013) and Singh et al. (2014) apply regular vines in the multivariate modeling to study the financial risk and return dependence between capital markets and stocks. Geidosch and Fischer (2016), in turn, demonstrate the superiority of vine copulas over conventional copulas when modeling the dependence structure of a credit portfolio. One of the authors' findings is that when mixing different copula families in an R-vine structure, the best statistical fit to the data can be achieved, which corresponds to the most reliable estimate for economic capital. This is in line with Singh and Allen (2017, 248–249), who show that the regular vine copula model provides a more flexible tree structure than the C- and D-vine copulas and can also give better results than the usual bivariate copula approach. On the other hand, Zhang et al. (2014) utilize regular, canonical, and drawable vines to forecast both the VaR and CVaR risk measures and conclude that the D-vine might be better than the other vine copulas in terms of forecasting CVaR.

An extensive study of canonical vine copulas in the context of portfolio management is also provided by Low et al. (2018). The authors utilize multi-dimensional elliptical and asymmetric copula models to manage downside correlations and forecast returns for portfolios with 3–12 US industry index constituents. Their analysis further assumes that investors have a utility function characterized by the minimization of CVaR. Low et al. (2018, 263, 287–288) find that for portfolios of higher dimensions, controlling for lower-tail dependence using the Clayton Canonical Vine Copula (CVC) and asymmetries within the skewed Student t marginals, the portfolio can insulate downside risk and, to some extent, protect the value of the portfolio with little loss to upside return. However, there seems to be little or no benefit to be gained by using a complex model of the dependence structure and marginals for simpler, smaller portfolios. In cases like these, using advanced models induces estimation error which overruns any benefits from the modeling, resulting in poor portfolio decisions. (see also Hatherley & Alcock 2007 for a similar conclusion.)

As a summary, Figure 2.4 illustrates the utilization of the R-vine copula in forecasting the Expected Shortfall for an equally-weighted portfolio consisting of nine sectoral price indices. The data was downloaded from the Moscow Exchange website and covers the period from 9 January 2008 to 7 September 2017 (2407 daily observations in total). Here, only the last 1000 log-returns are used, and the moving window is set to 250 days. The main steps can be outlined as follows: First, the GARCH(1,1) model with Student-t innovations is fitted to convert the log-returns into an i.i.d. series. Then, the extracted and standardized residuals are converted to Student-t marginals for the copula estimation. Once a multivariate matrix of uniform marginals has been obtained, an R-vine is fitted with the bivariate copula family consisting of six copulas (namely Gaussian, Student-t, Clayton, Gumbel, Frank, and Joe copula), and AIC is used as the selection criterion. Next, 1000 simulations per index are generated, and the simulated uniform marginals are converted to standardized residuals. Respectively, returns are simulated from the standardized residuals. Finally, a series of simulated daily portfolio returns is generated to forecast one-day-ahead ES with 99 and 95 percent confidence levels. (Singh & Allen 2017, 249–255; see also Allen et al. 2013; Nagler et al. 2019 for more technical details.)

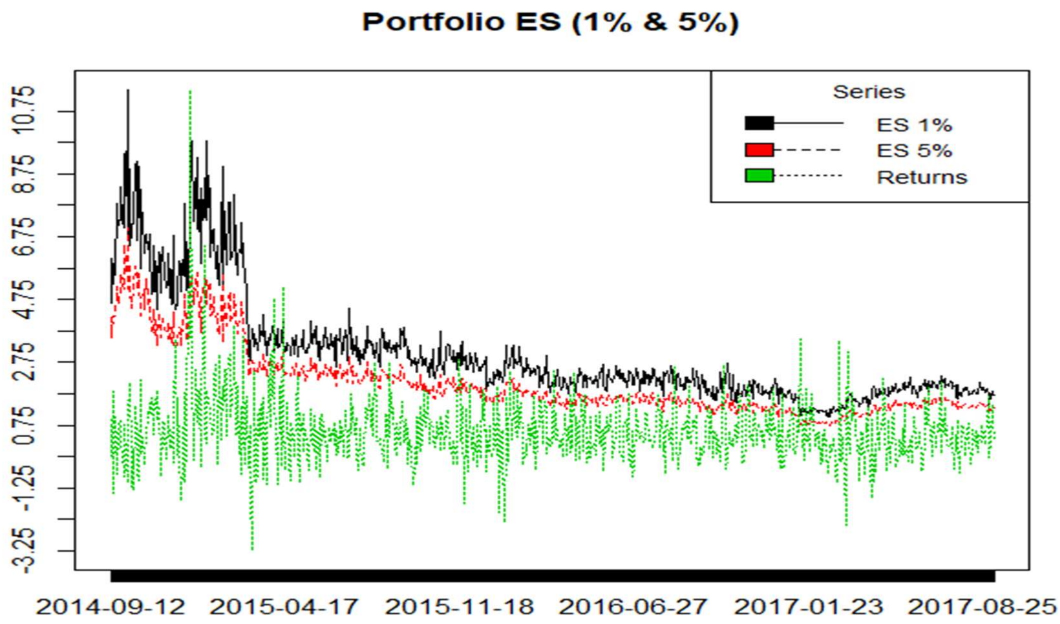

Figure 2.4 One and five percent R-vine ES forecasts along with an equally-weighted sectoral indices portfolio returns

The above approach leads to Expected Shortfall forecasts that are not time-dependent but have the advantage of being co-dependent on the assets in the portfolio. As can be seen, these forecasts closely follow the daily returns with few violations. A mere eyeball examination tells us that there are roughly ten exceedances at the 99 percent confidence level. Note that the model could be easily improved, e.g., by adding more copulas to the bivariate copula family. For example, Patton (2013, 947–952) demonstrates the use of time-varying copula models in portfolio risk management. His work is mainly based on the paper by Creal, Koopman, and Lucas (2012), who propose a class of observation-driven time series models referred to as generalized autoregressive score (GAS) models. Respectively, Almeida, Czado, and Manner (2016) promote the use of a dynamic D-vine model, which extends the stochastic autoregressive copula (SCAR) model by Hafner and Manner (2012), in the modeling of high-dimensional time-varying dependence and forecasting Value-at-Risk for daily stock returns.

3 DATA AND RESEARCH METHODS

We begin this section by providing a few descriptive statistics of the data used. Then, we discuss the principles of the algorithms used in performing different asset allocation strategies. The empirical findings are presented in the next chapter.

3.1 Data and descriptive statistics

In this thesis, we utilize two data sets representing the Russian stock market, namely daily sectoral price[1] indices downloaded from the Moscow Exchange (MOEX), as well as daily total return indices of 40 constituents of the MICEX index downloaded from the Thomson Reuters Datastream database. In both cases, the values are denominated in Russian rubles.

Let us start the depiction of the data with the nine sectoral indices. These include the Consumer Goods & Services (CGS), Chemicals (CHM), Electric Utilities (PWR), Financials (FNL), Manufacturing (MNF), Metals & Mining (M&M), Oil & Gas (O&G), Telecoms (TLC), and Transport (TRN) Index. As a benchmark, we use the MICEX price index. Our sample starts on December 28, 2007, and ends on September 7, 2017 (i.e., 2408 daily price observations in total). MOEX uses a free-float capitalization weighting approach, and the sector indices are calculated based on the prices (in RUB) of the most liquid shares of Russian issues admitted to trading in the public joint-stock company (PJSC) Moscow Exchange. These price return indices are calculated without considering the cash dividends on index securities. (see Moscow Exchange Sectoral Indices 2017.)

Table 3.1 lists the descriptive statistics. These include, inter alia, the minimum and maximum daily returns over the sample period, average annualized returns, annualized standard deviation, as well as skewness and kurtosis. Based on these metrics, we can observe that the Chemicals index has the most extreme range in returns (i.e., its minimum and maximum returns vary from -0.30 to 0.29), whereas the Consumer Goods & Services index has the smallest range with minimum and maximum returns varying from -0.14 to 0.10. Moreover, the Chemicals index has the highest annualized return of 0.11, whereas the Manufacturing index has the lowest annualized return of -0.08. In terms of risk, the Oil & Gas index has the highest annualized standard deviation of 0.34, whereas the Consumer Goods & Services index presents the lowest value of 0.22. Interestingly, the Oil & Gas index is also the only asset exhibiting positive skewness (0.59), whereas the Electric Utilities index has the most negative skewness of -1.09. Finally, the Chemicals (Consumer Goods & Services) index has the highest (lowest) kurtosis of 37.36 (12.34).

[1] When data was downloaded, total returns were not available for the sectoral indices.

In Table 3.1, we also report the results of the Jarque-Bera Test. The JB test statistic has an asymptotic χ^2-distribution with two degrees of freedom and can be used to test the null hypothesis that the data are from a normal distribution. The null hypothesis is a joint hypothesis of the sample skewness being zero and the excess kurtosis being zero. Any deviation from this increases the JB statistic. (see Jarque & Bera 1980; Bera & Jarque 1981; Cromwell, Labys & Terraza 1994, 20–22.) As can be deduced from the large values of the test statistic, all the indices depart from normality. This departure is also statistically significant since the p-values are smaller than 0.0001 for each index. Hence, we can reject the null hypothesis and conclude that the returns do not follow the normal distribution.

Next, we compute the sample autocorrelation coefficients for each time series at the first two lags (days). The statistically significant values at the five percent level are denoted with an asterisk. With respect to serial correlation, it is clear that if a return series is independently distributed, then the autocorrelation coefficients, $\rho(k)$, should be zero for all time lags of the differenced series (Cromwell et al. 1994, 25). However, as we can see in Table 3.1, all the first-order autocorrelation coefficients are positive and significant, except in the case of MICEX (0.01) and O&G (0.01). The TRN index is the only asset with a negative (and statistically significant) first-order AC coefficient of (-0.10*). Correspondingly, the second-order AC coefficients are negative and statistically significant, except in the case of the CGS (0.04), M&M (0.00), MNF (0.00), and TRN (0.04) index.

Although not reported here, we also conducted the Augmented Dickey-Fuller test (see Trapletti, Hornik & LeBaron 2020, 2–3) for these ten indices and found that only in the case of the Oil & Gas index was the p-value of the first-order DF test statistic smaller than the five percent significance level. Thus, we can conclude that this sector's price series possesses mean reversion (i.e., the next price level will be proportional to the current price level). Furthermore, to investigate the long-range dependence or long-memory behavior of each time series process, we calculated the Hurst Exponents using both the simple R/S Hurst estimation and the differenced aggregated variance methods (see Hurst 1951; Weron 2002; Würtz et al. 2017a, 3; Borchers 2021, 167-169 for more technical details). We found that for each series, the Hurst Exponent was larger than 0.5 ($H > 0.5$), meaning that the time series of Russian sectoral indices are persistent (i.e., trending).

In Table 3.1, we consider the forecastability of Russian equities a bit more specifically. The best forecasting method for stock market indices (and prices) is often the *naïve method*. In other words, each forecast is simply equal to the last observed value, or $\hat{y}_t = y_{t-1}$. Hence, the residuals are equal to the difference between consecutive observations: $e_t = y_t - \hat{y}_t = y_t - y_{t-1}$. We obtain the residuals from forecasting each individual time series using this naïve method with daily closing observations (i.e., levels). Under the efficient market hypothesis, the naïve method should produce forecasts that account for all available information, meaning, inter alia, that no significant correlation exists in the residuals series. In order to test jointly that several autocorrelations are zero, we compute the Ljung-Box test statistic (see Ljung & Box 1978; Cromwell et al. 1994, 25–27; Tsay 2010, 32–35; Halls-Moore 2017, 125) for the null hypothesis $H_0 : \rho_1 = \cdots = \rho_h = 0$ against the alternative hypothesis $H_a : \rho_i \neq 0$ for some $i \in \{1, \ldots, h\}$. In other words, we examine whether the first h autocorrelations are significantly different from what would be expected from a white noise process. In line with Hyndman and Athanasopoulos (2014, 52–57), we set the maximum lag being considered to ten ($h = 10$). If the autocorrelations did come from a white noise series, then the test statistic Q^* would have a χ^2-distribution with h degrees of freedom.

We observe that the LB $Q^*(10)$ statistic for the cumulative effect of up to tenth-order autocorrelation in the residuals is large, with p-values smaller than the five percent significance level. Hence, we can reject the zero autocorrelation null hypothesis for each index and conclude that the residuals are not a realization of discrete white noise, meaning that the time series of daily returns is not independent. Since the distribution of the next return seems to depend not only on the current return but on several previous returns, our sample has all of the well-documented characteristics of the unconditional distribution of equity returns. These stylized facts can then be used to justify the use of various (parametric) forecasting models. Indeed, if the variance is a risk proxy, then higher variance may lead to higher levels of expected return. In order to capture the linear dependence in the residuals, one could employ, say, an Autoregressive Moving Average (ARMA) model for the conditional mean in a time-series model for conditional variance.

Finally, we conduct Welch's two-sample t-test (see Welch 1947) in order to analyze whether the true difference in the average (mean) returns between a sectoral index and the MICEX Index is statistically equal to zero. Note that this test is more robust than Student's t-test when the two samples have unequal variances. Judging by the high two-tailed p-values, which are all much greater than the five percent significance level, we cannot reject the null hypothesis that the two population means are equal (i.e., the difference is not statistically significant). This conclusion is not hugely surprising, given that, despite the central limit theorem, we are using return data that violates the assumption of normality.

Table 3.1 Descriptive statistics of the price indices returns of the MICEX index and the sectoral indices based on daily observations

							Jarque-Bera test for normality				*Ljung-Box test for the residuals from the naïve method with 10 lags*		*Welch two-sample t-test*	
	Min	Max	Ann. return	Ann. Std Dev	Skewness	Kurtosis	χ^2	p-value	ACF (lag 1)	ACF (lag 2)	Q*	p-value	t-statistic	p-value
MICEX	-0.207	0.252	0.007	0.333	-0.074	24.682	61102	0.000	0.014	-0.045*	626	0.000	-	-
CGS	-0.135	0.103	0.074	0.215	-0.821	12.340	15542	0.000	0.114*	0.035	664	0.000	0.524	0.601
CHM	-0.301	0.285	0.111	0.333	-0.539	37.355	140064	0.000	0.096*	-0.043*	847	0.000	0.687	0.492
FNL	-0.231	0.254	0.005	0.337	-0.253	22.601	51257	0.000	0.041*	-0.068*	588	0.000	-0.010	0.992
M&M	-0.202	0.177	0.040	0.337	-0.624	14.492	21219	0.000	0.131*	0.001	916	0.000	0.217	0.828
MNF	-0.152	0.183	-0.080	0.317	-0.646	13.940	19656	0.000	0.084*	0.002	847	0.000	-0.578	0.563
O&G	-0.198	0.299	0.057	0.342	0.594	28.340	80692	0.000	0.013	-0.050*	622	0.000	0.324	0.746
PWR	-0.252	0.217	-0.061	0.327	-1.086	22.774	52492	0.000	0.138*	-0.043*	915	0.000	-0.444	0.657
TLC	-0.203	0.175	-0.043	0.303	-1.032	17.162	29967	0.000	0.094*	-0.065*	769	0.000	-0.337	0.736
TRN	-0.194	0.191	-0.006	0.327	-0.072	15.373	23704	0.000	-0.098*	0.037	442	0.000	-0.084	0.933

It is often interesting to see how the return series move in relation to each other and how much they are related. Thus, Table 3.2 shows a matrix of Pearson's correlation coefficients for the MICEX and sectoral price indices. Unsurprisingly, the highest correlation is between the Oil & Gas and the MICEX index (0.97). The O&G index is also highly correlated with the other sectoral indices, the highest coefficients being with the Metals and Mining (0.82), as well as with the Financials (0.81) index. In contrast, the Transport sector is the least correlated with the MICEX index (0.49), as well as with the other sectoral indices (values ranging from 0.41 to 0.49). All these correlation coefficients are statistically significant at the five percent level (i.e., $p < 0.0001$), and hence they have been marked with an asterisk. Furthermore, although not reported here, it is worth mentioning that the rolling 252-day correlations between the MICEX index and the sectoral indices peaked in February–March 2014 when the Ukrainian crisis started and decreased equally rapidly to the pre-crisis levels at the beginning of 2015.

Table 3.2 Correlation matrix for the MICEX and sectoral price indices

	MICEX	*CGS*	*CHM*	*FNL*	*M&M*	*MNF*	*O&G*	*PWR*	*TLC*	*TRN*
MICEX	**1.000**									
CGS	0.606*	**1.000**								
CHM	0.710*	0.505*	**1.000**							
FNL	0.875*	0.574*	0.625*	**1.000**						
M&M	0.857*	0.576*	0.656*	0.712*	**1.000**					
MNF	0.590*	0.571*	0.519*	0.585*	0.566*	**1.000**				
O&G	0.972*	0.571*	0.691*	0.814*	0.816*	0.560*	**1.000**			
PWR	0.732*	0.608*	0.581*	0.694*	0.670*	0.591*	0.686*	**1.000**		
TLC	0.795*	0.607*	0.590*	0.741*	0.709*	0.572*	0.758*	0.694*	**1.000**	
TRN	0.495*	0.415*	0.408*	0.478*	0.428*	0.408*	0.469*	0.479*	0.463*	**1.000**

Figure 3.1 presents the daily log-returns of the MICEX price index as a bar chart overlayed with two rolling measures of tail risk, namely the traditional Gaussian Value-at-Risk and more conservative Expected Shortfall at the 95 percent confidence level (see Peterson et al. 2015, 97–98; Peterson et al. 2018, 11). The rolling 1-day risk calculations help us identify events and periods when estimates of tail risk changed suddenly. They also help us evaluate whether the assumptions underlying the calculation seem to hold. Overall, based on these dynamic estimates, the risk level has been steadily decreasing in the Russian stock market since the global financial crisis of 2008. On the other hand, the outbreak of the Russo-Ukrainian War at the beginning of 2014 caused a sudden and unexpected negative peak in volatility.

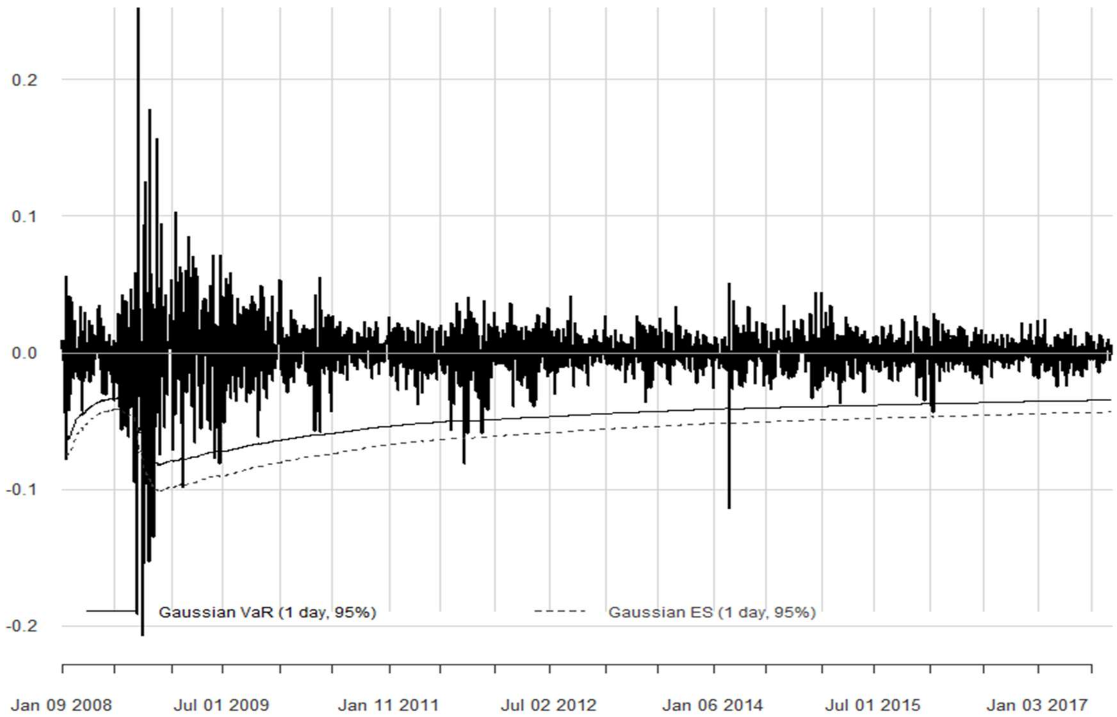

Figure 3.1 Daily returns of the MICEX price index with risk metric overlay

Let us now focus on the individual stocks. Unfortunately, since the Russian stock market is relatively young and the amount of data is limited, we need to choose between the length of time series and the number of investable equities. Therefore, as a compromise, we decide to use the daily total return indices data of 40 (out of roughly 50) constituent stocks of the MICEX gross Total Return (TR) index from 21 July 2011 to 6 September 2017. As such, our sample period encompasses over six years, consisting of 1538 daily observations in total. We remove firms that do not have information on stock returns available for the whole period. The total return indices reflect cum-dividend prices and account for all further corporate actions, reinvested dividends, and stock splits, making them the most suitable metric for return calculations (Fischer & Krauss 2018, 655). For more information on the computation methodology of these indices, see Thomson Reuters Global Equity Indices (2015, 8–10). The MICEX gross TR index, in turn, is calculated by the Moscow Exchange. It is designed to represent the overall return of its constituents (i.e., their capital performance and income from reinvesting dividends). Here, 'gross' means that the TR index is calculated subject to gross dividends (i.e., dividend income taxes are not considered). (see Total Return Indices 2017.)

Table 3.3 shows the company names and tickers of these 40 ordinary and preferred shares representing nine different industrial sectors. The additional two assets at the bottom of the list, namely Far-Eastern Shipping Company (FESH) and TransContainer (TRCN), represent the stocks from the Transport sector used in our copula opinion pooling application. Overall, these are the most liquid stocks of the Moscow Exchange. Furthermore, our sample should be free of survivorship bias since none of the constituents of the MICEX index has been delisted from the stock exchange during the period considered. The reader interested in learning more about the features of the Russian preferred shares is referred, e.g., to Liljeblom and Maury (2016).[2]

The descriptive statistics of these assets are presented in Table 3.4. The minimum returns vary from -0.53 (Mechel) to -0.08 (Rosneft), whereas the maximum returns vary from 0.05 (the MICEX index) and 0.07 (Surgutneftegas) to 0.83 (PhosAgro). AFK Sistema has the largest return range from -0.46 and 0.72. Furthermore, Mechel (PhosAgro) has the lowest (highest) average annualized return of -0.23 (0.31). In terms of annualized standard deviation, Mechel is also the riskiest asset (0.69), whereas the MICEX index (0.20) and Lukoil (0.25) are at the opposite end of the spectrum.

As is common with stocks, most of our assets (27 out of 42, i.e., 64 percent) exhibit negative skewness, implying more negative extreme returns. Alrosa has the most negative skewness of -2.11, whereas PhosAgro has the most positive skewness of 15.89. Similarly, most of the return distributions (32 out of 42, i.e., 76 percent) exhibit noticeable excess

[2] In brief, a preferred stock may provide some shareholder protection because its dividend is determined as a fixed percentage of net profits. Thus, it is more difficult to channel cash from the firm to the controlling owner proportionally to voting rights. (Liljeblom & Maury 2016, 2416–2417.)

kurtosis. Again, PhosAgro has the highest kurtosis of 461.10, whereas Tatneft has the lowest kurtosis of 1.13. The presence of leptokurtosis means that large moves occur far more frequently than one would expect if the returns were normally distributed (Sinclair 2013, 40). Hence, a more sophisticated approach than standard deviation (or volatility) is required to model the risk adequately.

Furthermore, the Jarque-Bera test rejects the null hypothesis of normality for each of the 43 assets at the five percent significance level. Specifically, the values of the JB test statistics are large and range from 82 (Tatneft) to 13,680,893 (PhosAgro). Correspondingly, the test's p-values are all highly significant (i.e., $p < 0.0001$). Hence, we can conclude that none of the return series is normally distributed, which is in line with the well-known stylized facts for financial market returns.

Considering the linear dependence between the time series and its past values, we compute the autocorrelation functions (ACFs) at the first two lags. We can see that 19 assets out of 43 (44 percent) have a statistically significant first-order autocorrelation coefficient at the five percent level (denoted with an asterisk). Out of these nineteen assets, only two have a negative autocorrelation coefficient, namely Unipro (-0.07) and TransContainer (-0.17). Mechel has the highest positive first-order autocorrelation coefficient of 0.14. With respect to the second-order ACFs, ten out of 43 assets (23 percent) have a statistically significant autocorrelation coefficient, and only one of these, namely TMK, has a significant positive coefficient of 0.07. Alrosa has the most negative second-order AC coefficient of -0.14. Thus, there seems to be momentum at the first lag, which, to a lesser extent, is followed by mean-reversion at the second lag. Naturally, when autocorrelation is high, it becomes easier to predict future values by referring to past values. To put it another way, when data have a trend, the autocorrelations for small lags tend to be large and positive because observations nearby in time are also nearby in size.[3] (Tsay 2010, 31; Hyndman & Athanasopoulos 2014, 35.)

Based on the p-values of the Augmented Dickey-Fuller test statistic with a lag order value of one (not reported here), eight of these 42 stocks are statistically significantly mean-reverting at the five percent level. These include Gazprom, Lukoil, LSR Group, Mechel, Rostelecom, Tatneft, Unipro, and Uralkali. It is worth noting that since three of these companies belong to the same Oil & Gas sector, one could exploit the tendency of these time series to revert to a historical mean value by creating, say, a classic 'pairs trade' trading strategy (see, e.g., Halls-Moore 2015, 172–180). We leave such considerations for future research. On the other hand, the Hurst Exponents (not reported here) for each time series are near one, indicating that they are strongly trending. This finding shows again that the daily prices of Russian equities do not behave akin to geometric Brownian

[3] However, it is worth recalling that autocorrelation measures linear relationships; even if the autocorrelation is minuscule, there can still be a nonlinear relationship between a time series and a lagged version of itself.

motion (i.e., long-term positive autocorrelation is a distinctive feature). Therefore, we should be able to capitalize on this behavior by forming momentum trading strategies.

The Ljung-Box test statistic, $Q^*(10)$, on the residuals from the naïve method provides us with another formal test of intertemporal dependence in the variance. As we can see in Table 3.4, the values of this statistic are large, with p-values less than the five percent significance level (i.e., $p < 0.0001$). As the results are statistically significant, we can reject the null hypothesis and conclude that the residuals are distinguishable from a white noise series. Unlike in more efficient Western stock markets (see, e.g., Tsay 2010, 34; Hyndman & Athanasopoulos 2014, 216), the daily change in the Russian stock prices is correlated with that of previous days and thus not completely random. This suggests that we should be able to construct deterministic models and compute forecasts that utilize the information left in the residuals. Indeed, the fact that the residuals are not behaving like white noise is in contrast to the 'random walk' model, stating that future movements (i.e., daily changes in direction) are unpredictable. Note that the large values of Q^* for certain assets (e.g., 748 of Rosseti) are most likely a result of the nonsynchronous trading effect (see Kim & Kon 1994, 566).

With respect to Welch's two-sample t-test, we can see that the two-tailed p-values (ranging from 0.15 of Rostelecom to 0.99 of Bashneft, pref.) are greater than the five percent significance level, meaning that we cannot reject the null hypothesis of equal population means. However, although the difference is not statistically significant, 24 out of 40 stocks (60 percent) have a positive t-statistic. Hence, it might be possible to construct portfolios that yield higher mean returns than the MICEX index. It can also be concluded that the gap between the best and worst-performing assets is large. For example, Tatneft pref. (1.26), PhosAgro (1.20), Surgutneftegas pref. (1.16), and Acron Group (1.06) have the highest test statistic, whereas the opposite is true for Rostelecom (-1.45), Rosseti (-1.31), Uralkali (-1.09), and Mechel (-1.04). Finally, based on a static view of the relationships over time (not reported here), Polyus and Alrosa are the least correlated with the MICEX index (0.19 and 0.26, respectively), whereas Gazprom and Sberbank have the highest correlation with the MICEX index (0.85 and 0.80, respectively). All these correlation coefficients are statistically significant at the five percent level.

Table 3.3 Asset universe consisting of 40 constituent stocks of the MICEX gross total return index

	Company name	*Ticker*	*Sector*
1	Acron Group (Akron)	AKRN	Chemicals
2	PhosAgro (FosAgro)	PHOR	Chemicals
3	Uralkali (Uralkaliy)	URKA	Chemicals
4	Dixy Group	DIXY	Consumer G&S
5	Magnit	MGNT	Consumer G&S
6	M.video	MVID	Consumer G&S
7	Federal Grid Company of Unified Energy System (FGC UES)	FEES	Electric Utilities
8	Inter RAO UES	IRAO	Electric Utilities
9	Mosenergo	MSNG	Electric Utilities
10	Rosseti (Rossiyskiye Seti)	RSTI	Electric Utilities
11	RusHydro	HYDR	Electric Utilities
12	Unipro (Yunipro)	UPRO	Electric Utilities
13	AFK Sistema	AFKS	Financials
14	Bank VTB	VTBR	Financials
15	Sberbank of Russia (Sberbank Rossii)	SBER	Financials
16	Sberbank of Russia, Pref.	SBERP	Financials
17	Alrosa	ALRS	Metals & Mining
18	Magnitogorsk Iostl. Works (Magnitogorski metallurgitšeski kombinat)	MAGN	Metals & Mining
19	Mechel (Metšel)	MTLR	Metals & Mining
20	Norilsk Nickel (Norilski nikel)	GMKN	Metals & Mining
21	Novolipetsk Steel (Novolipetski metallurgitšeski kombinat)	NLMK	Metals & Mining
22	Polyus	PLZL	Metals & Mining
23	Severstal	CHMF	Metals & Mining
24	TMK (Trubnaya Metallurgicheskaya Kompaniya)	TRMK	Metals & Mining
25	Bashneft, Pref.	BANEP	Oil & Gas
26	Gazprom	GAZP	Oil & Gas
27	Lukoil	LKOH	Oil & Gas
28	Novatek	NVTK	Oil & Gas
29	Rosneft Oil Co	ROSN	Oil & Gas
30	Surgutneftegas	SNGS	Oil & Gas
31	Surgutneftegas, Pref.	SNGSP	Oil & Gas
32	Tatneft	TATN	Oil & Gas
33	Tatneft, Pref.	TATNP	Oil & Gas
34	Transneft, Pref.	TRNFP	Oil & Gas
35	LSR Group	LSRG	Real Estate
36	PIK Group	PIKK	Real Estate
37	Mobile TeleSystems (Mobil'nye Telesistemy)	MTSS	Telecoms
38	Rostelecom (Rostelekom)	RTKM	Telecoms
39	Aeroflot Russ. Airl. (Aeroflot - Rossiyskiye Avialinii)	AFLT	Transport
40	Novorossiysk Commercial Sea Port (Novorossiyskiy morskoy torgovyi port)	NMTP	Transport

Table 3.4 Descriptive statistics of the total returns of the MICEX index and its 40 constituents based on daily observations

							Jarque-Bera test for normality				*Ljung-Box test for the residuals from the naïve method with 10 lags*		*Welch two-sample t-test*	
	Min	Max	Ann. return	Ann. Std Dev	Skew-ness	Kurto-sis	χ^2	p-value	ACF (lag 1)	ACF (lag 2)	Q*	p-value	t-statistic	p-value
MICEX	-0.114	0.051	0.072	0.204	-0.787	6.289	2691	0.000	0.033	-0.027	432	0.000	-	-
AFKS	-0.461	0.724	-0.098	0.581	1.749	135.264	1172512	0.000	0.106*	-0.077*	218	0.000	-0.680	0.497
AFLT	-0.182	0.103	0.203	0.353	-0.631	7.070	3303	0.000	0.091*	-0.005	374	0.000	0.795	0.427
AKRN	-0.140	0.069	0.225	0.296	-0.515	4.149	1170	0.000	0.028	0.025	249	0.000	1.057	0.291
ALRS	-0.511	0.337	0.200	0.434	-2.112	111.058	791030	0.000	0.104*	-0.136*	341	0.000	0.663	0.507
BANEP	-0.221	0.192	0.074	0.343	-0.535	15.632	15723	0.000	0.064*	0.009	490	0.000	0.014	0.989
CHMF	-0.115	0.088	0.169	0.337	-0.124	1.952	248	0.000	0.059*	-0.010	286	0.000	0.609	0.542
DIXY	-0.212	0.169	-0.039	0.346	-0.468	12.006	9287	0.000	0.120*	-0.009	465	0.000	-0.679	0.497
FEES	-0.247	0.156	-0.100	0.454	-0.506	7.462	3631	0.000	0.095*	-0.010	526	0.000	-0.853	0.394
GAZP	-0.150	0.076	-0.038	0.259	-0.343	5.800	2185	0.000	0.036	-0.004	390	0.000	-0.822	0.411
GMKN	-0.159	0.121	0.132	0.288	-0.273	6.577	2789	0.000	0.033	-0.061*	293	0.000	0.421	0.674
HYDR	-0.099	0.125	-0.048	0.345	0.212	3.057	610	0.000	0.035	-0.023	390	0.000	-0.739	0.460
IRAO	-0.183	0.115	0.009	0.388	-0.031	4.490	1291	0.000	0.103*	0.036	506	0.000	-0.354	0.723
LKOH	-0.103	0.092	0.128	0.248	-0.182	3.069	612	0.000	0.014	-0.063*	386	0.000	0.438	0.662
LSRG	-0.200	0.085	0.038	0.344	-0.639	7.083	3317	0.000	0.022	-0.001	374	0.000	-0.205	0.838
MAGN	-0.164	0.095	0.111	0.367	-0.211	3.428	764	0.000	0.056*	0.031	315	0.000	0.231	0.817
MGNT	-0.123	0.099	0.183	0.331	-0.200	2.700	477	0.000	-0.016	-0.031	339	0.000	0.704	0.481
MSNG	-0.160	0.157	0.064	0.334	0.264	6.992	3149	0.000	0.075*	0.039	496	0.000	-0.046	0.963
MTLR	-0.534	0.448	-0.232	0.694	-0.454	30.558	59856	0.000	0.144*	-0.062*	452	0.000	-1.036	0.301
MTSS	-0.142	0.225	0.106	0.298	0.423	17.692	20090	0.000	-0.015	-0.014	271	0.000	0.235	0.814
MVID	-0.169	0.094	0.168	0.327	-0.734	8.422	4681	0.000	0.105*	-0.006	382	0.000	0.616	0.538
NLMK	-0.133	0.105	0.082	0.349	-0.037	2.650	450	0.000	0.058*	0.004	376	0.000	0.063	0.950
NMTP	-0.174	0.176	0.212	0.361	0.339	8.834	5027	0.000	0.044	-0.003	366	0.000	0.840	0.401
NVTK	-0.170	0.094	0.084	0.311	-0.538	6.060	2426	0.000	-0.042	-0.031	258	0.000	0.085	0.932
PHOR	-0.120	0.834	0.314	0.455	15.891	461.102	13680893	0.000	0.018	-0.018	361	0.000	1.199	0.231
PIKK	-0.091	0.114	0.182	0.293	0.474	4.186	1180	0.000	0.036	0.037	337	0.000	0.761	0.447
PLZL	-0.136	0.300	0.187	0.363	2.758	28.238	53013	0.000	0.132*	-0.019	396	0.000	0.688	0.492
ROSN	-0.079	0.071	0.075	0.264	0.082	1.326	114	0.000	0.063*	-0.007	453	0.000	0.024	0.981
RSTI	-0.259	0.162	-0.203	0.477	-0.291	7.438	3565	0.000	0.080*	-0.021	748	0.000	-1.309	0.191
RTKM	-0.158	0.131	-0.145	0.309	-0.016	8.090	4192	0.000	0.005	-0.064*	379	0.000	-1.448	0.148
SBER	-0.161	0.120	0.118	0.331	-0.417	5.399	1911	0.000	0.049	-0.078*	401	0.000	0.297	0.767
SBERP	-0.132	0.102	0.141	0.330	-0.423	3.407	789	0.000	0.047	-0.029	433	0.000	0.442	0.659
SNGS	-0.109	0.067	0.019	0.285	0.052	1.912	235	0.000	0.027	-0.031	381	0.000	-0.368	0.713
SNGSP	-0.095	0.084	0.234	0.280	-0.166	2.435	387	0.000	0.030	-0.004	473	0.000	1.161	0.246
TATN	-0.092	0.082	0.159	0.341	0.043	1.127	82	0.000	-0.026	-0.048	287	0.000	0.544	0.586
TATNP	-0.089	0.112	0.237	0.251	0.113	3.764	911	0.000	0.004	-0.023	268	0.000	1.263	0.207
TRMK	-0.310	0.290	-0.067	0.408	0.073	32.308	66847	0.000	-0.015	0.068*	193	0.000	-0.748	0.454
TRNFP	-0.156	0.147	0.240	0.358	0.027	5.888	2221	0.000	0.045	-0.054*	299	0.000	1.011	0.312
UPRO	-0.178	0.084	0.066	0.310	-0.519	6.643	2895	0.000	-0.072*	-0.011	213	0.000	-0.035	0.972
URKA	-0.233	0.136	-0.098	0.325	-1.661	22.139	32097	0.000	0.109*	-0.032	451	0.000	-1.092	0.275
VTBR	-0.193	0.101	-0.042	0.303	-0.525	8.810	5042	0.000	0.051	-0.007	400	0.000	-0.769	0.442
Additional stocks of the Transport sector used in the Copula Opinion Pooling application:														
FESH	-0.136	0.263	-0.168	0.449	1.814	14.322	13979	0.000	-0.031	-0.080*	98	0.000	-1.200	0.230
TRCN	-0.285	0.200	0.064	0.428	-0.488	18.293	21492	0.000	-0.165*	-0.039	172	0.000	-0.042	0.967

3.2 Classical and robust mean-risk portfolios

We start this subchapter by constructing four optimal portfolios based on ordinary sample estimators with the aim to outperform the MICEX Total Return Gross Index (thus, our investable universe consists of the 40 component stocks of this benchmark). After that, we move on to the robust covariance estimation and robust optimization techniques. Within the mean-risk framework, the first obvious choice is to consider a variance minimization strategy, which is a quadratic problem of the form:

$$\min_{\boldsymbol{\omega}} \boldsymbol{\omega}'\Sigma\boldsymbol{\omega}, \tag{3.2.1}$$

where Σ is the estimated covariance matrix of asset returns and $\boldsymbol{\omega}$ is the set of weights. The second option is the return maximation strategy, which lies on the other end of the Markowitz efficient frontier. This linear problem can be defined as follows:

$$\max_{\boldsymbol{\omega}} \hat{\boldsymbol{\mu}}'\boldsymbol{\omega}, \tag{3.2.2}$$

where $\hat{\boldsymbol{\mu}}$ is the estimated mean asset returns.[4] (see Bennett 2015a, 5–6; 2015b, 3–14 for more technical details about the R algorithms used.) For the sake of experimentation, we consider two lesser-known modifications of this latter strategy.

Indeed, as our third strategy, we maximize the mean return per unit Expected Shortfall (Conditional Value-at-Risk or Expected Tail Loss). This optimization problem is based on the coherent reward-risk ratio called *Stable Tail-Adjusted Return Ratio* (STARR).[5] Following the notation used by Stoyanov et al. (2007, 415) and Rachev et al. (2008, 31), the maximization of the STARR ratio can be formally defined as:

$$\max_{\boldsymbol{\omega}} STARR(\boldsymbol{\omega}) = \max_{\boldsymbol{\omega}} \frac{E(\boldsymbol{\omega}'\boldsymbol{r} - r_f)}{ES_\alpha(\boldsymbol{\omega}'\boldsymbol{r} - r_f)}, \tag{3.2.3}$$

where the vector notation $\boldsymbol{\omega}'\boldsymbol{r}$ stands for the returns of a portfolio with composition $\boldsymbol{\omega} = (\omega_1, \omega_2, \dots, \omega_n)$, and E is the expectation operator. Here, the confidence level is set to 95 percent ($\alpha = 0.05$), and the risk-free interest rate (r_f) is set to zero percent for simplicity.

The fourth portfolio provides a demonstration of the risk budget mean-ES optimization. In other words, it seeks to maximize mean return per unit ES (CVaR or ETL) with ES equal contribution to risk. This approach is proposed by Boudt, Peterson, and Croux (2008), who provide a comprehensive exposition on the computation and decomposition of portfolio risk. As a brief illustration, consider an algorithm that allocates the portfolio between n assets, with weights $\boldsymbol{\omega} = (\omega_1, \dots, \omega_n)'$. The n asset returns are stacked into the random vector $\boldsymbol{r} = (r_1, \dots r_n)'$, which is assumed to be strictly stationary with mean $\boldsymbol{\mu}$ and covariance matrix Σ. It follows that the portfolio return $\boldsymbol{r}_p$ has mean $\boldsymbol{\omega}'\boldsymbol{\mu}$ and variance

[4] Alternatively, one could follow Estrada (2010), as well as De Santiago and Estrada (2013), who maximize the geometric mean return.

[5] See, e.g., Martin, Rachev and Siboulet (2003), Biglova et al. (2004), Rachev et al. (2007), Stoyanov, Rachev and Fabozzi (2007), Rachev et al. (2008), Chia, Lim and Chan (2013), Goel, Sharma and Mehra (2019) for further discussions.

$\boldsymbol{\omega}'\Sigma\boldsymbol{\omega}$. Under the additional assumption that the portfolio return distribution $F(\cdot)$ is continuous, the VaR and ES of $\boldsymbol{r}_p$ as a function of a loss probability α, can be defined as:

$$\begin{aligned} \text{VaR}(\alpha) &= -F^{-1}(\alpha) \\ \text{ES}(\alpha) &= -E_F\left[\boldsymbol{r}_p | \boldsymbol{r}_p \leq F^{-1}(\alpha)\right], \end{aligned} \tag{3.2.4}$$

with $F^{-1}(\cdot)$ the quantile function associated to $F(\cdot)$ and $E_F[\cdot \mid \cdot]$ the operator that takes the conditional expectation under $F(\cdot)$.

As shown in Boudt et al. (2008), the risk measures are required to be 1-homogeneous for portfolio risk decomposition. That is, if the weight vector is multiplied by some scalar b, then also these risk measures are multiplied by b. One can now employ Euler's homogeneous function theorem, as it states that for 1-homogeneous $f(\boldsymbol{\omega})$, there is

$$f(\boldsymbol{\omega}) = \sum_{i=1}^{n} \omega_i \partial_i f(\boldsymbol{\omega}), \tag{3.2.5}$$

where $\partial_i f(\boldsymbol{\omega}) = \partial f(\boldsymbol{\omega})/\partial\omega_i$. Under this decomposition, the Contribution of asset i to the risk measure $f(\boldsymbol{\omega})$, $C_i f(\boldsymbol{\omega})$, and its percentage Contribution, $\%C_i f(\boldsymbol{\omega})$, which is also known as the (percentage) Component of asset i in the portfolio risk measure $f(\boldsymbol{\omega})$, equal

$$C_i f(\boldsymbol{\omega}) = \omega_i \partial_i f(\boldsymbol{\omega}) \quad \text{and} \quad \%C_i f(\boldsymbol{\omega}) = C_i f(\boldsymbol{\omega})/f(\boldsymbol{\omega}). \tag{3.2.6}$$

We can interpret the contribution to $ES(\alpha)$ as the expected contribution to portfolio return when the portfolio return is at least the negative value of VaR(α):

$$C_i \text{ES}(\alpha) = \omega_i \partial_i \text{ES}(\alpha) = -E\left[\omega_i r_i | \boldsymbol{r}_p \leq -\text{VaR}(\alpha)\right]. \tag{3.2.7}$$

Since we are dealing with a downside risk measure that depends on the portfolio moments, it is computationally convenient to express the portfolio moments as a function of the multivariate moments of the returns on the underlying assets. This can be done by using the $N \times N^2$ co-skewness matrix $M_3 = E[(\boldsymbol{r}-\boldsymbol{\mu})(\boldsymbol{r}-\boldsymbol{\mu})' \otimes (\boldsymbol{r}-\boldsymbol{\mu})']$ and $N \times N^3$ co-kurtosis matrix $M_4 = E[(\boldsymbol{r}-\boldsymbol{\mu})(\boldsymbol{r}-\boldsymbol{\mu})' \otimes (\boldsymbol{r}-\boldsymbol{\mu})' \otimes (\boldsymbol{r}-\boldsymbol{\mu})']$, where $\otimes$ stands for the Kronecker product. Under this representation, the derivatives of the portfolio moments can be computed straightforwardly. If we denote the q-th centered portfolio moment $m_q = E\left[\left(\boldsymbol{r}_p - \boldsymbol{\omega}'\boldsymbol{\mu}\right)^q\right]$ and let $\partial_i m_q$ be its partial derivative with respect to the portfolio weight w_i, we get

$$\begin{aligned} m_2 &= \boldsymbol{\omega}'\Sigma\boldsymbol{\omega} & \partial_i m_2 &= 2(\Sigma\boldsymbol{\omega})_i \\ m_3 &= \boldsymbol{\omega}' M_3(\boldsymbol{\omega} \otimes \boldsymbol{\omega}) & \partial_i m_3 &= 3\big(M_3(\boldsymbol{\omega} \otimes \boldsymbol{\omega})\big)_i \\ m_4 &= \boldsymbol{\omega}' M_4(\boldsymbol{\omega} \otimes \boldsymbol{\omega} \otimes \boldsymbol{\omega}) & \partial_i m_4 &= 4\big(M_4(\boldsymbol{\omega} \otimes \boldsymbol{w} \otimes \boldsymbol{\omega})\big)_i. \end{aligned} \tag{3.2.8}$$

The portfolio's skewness s_p and excess kurtosis k_p and their partial derivatives are then given by

$$\begin{aligned} s_p &= m_3/m_2^{3/2} & \partial_i s_p &= \left(2m_2^{3/2}\partial_i m_3 - 3m_3 m_2^{1/2}\partial_i m_2\right)/2m_2^3 \\ k_p &= m_4/m_2^2 - 3 & \partial_i k_p &= (m_2 \partial_i m_4 - 2m_4 \partial_i m_2)/m_2^3. \end{aligned} \tag{3.2.9}$$

(see Boudt et al. 2008, 85; Pfaff 2016, 42–43.)

The computation of the derivative of the estimated ES is difficult for many estimation methods because the estimator cannot be expressed as an explicit function of the portfolio weights. Nevertheless, Boudt et al. (2008, Appendix C) show that there is a way to compute even the derivative of modified Expected Shortfall ($\partial_i \mathrm{mES}(\alpha)$) analytically. This approach, although tedious, is employed here. By definition, *modified ES* for a loss probability α is defined as the expected value of all returns below the α Cornish-Fisher quantile and where the expectation is computed under the second-order Edgeworth expansion of the true distribution function $G(\cdot)$:

$$\mathrm{mES}(\alpha) = -\boldsymbol{\omega}'\boldsymbol{\mu} - \sqrt{m_2} E_{G_2}[z|z \leq g_\alpha], \tag{3.2.10}$$

with $g_\alpha = G_2^{-1}(\alpha)$. In the above equation, $E_{G_2}[z|z \leq g_\alpha]$ can be expressed as a polynomial in g_α with coefficients that depend on the portfolio skewness s_p and excess kurtosis k_p, as well as on the standard Gaussian density function $\phi(\cdot)$. More precisely,

$$\begin{aligned} E_{G_2}[z|z \leq g_\alpha] = -\frac{1}{\alpha}\{\phi(g_\alpha) + \frac{1}{24}[I^4 - 6I^2 + 3\phi(g_\alpha)]k_p + \frac{1}{6}[I^3 - 3I^1]s_p \\ + \frac{1}{72}[I^6 - 15I^4 + 45I^2 - 15\phi(g_\alpha)]s_p^2\}, \end{aligned} \tag{3.2.11}$$

where

$$I^q = \begin{cases} \sum_{i=1}^{q/2} \left(\frac{\prod_{j=1}^{q/2} 2j}{\prod_{j=1}^{i} 2j}\right) g_\alpha^{2i}\phi(g_\alpha) + \left(\prod_{j=1}^{q/2} 2j\right)\phi(g_\alpha) & \text{for } q \text{ even} \\ \sum_{i=0}^{q^*} \left(\frac{\prod_{j=0}^{q^*}(2j+1)}{\prod_{j=0}^{i}(2j+1)}\right) g_\alpha^{2i+1}\phi(g_\alpha) - \left(\prod_{j=0}^{q^*}(2j+1)\right)\Phi(g_\alpha) & \text{for } q \text{ odd} \end{cases} \tag{3.2.12}$$

and $q^* = (q-1)/2$ and $\Phi(\cdot)$ is the standard Gaussian distribution function.

Finally, using the property $\phi'(z) = -z\phi(z)$, the derivative of modified ES can be computed as follows:

$$\begin{aligned} \partial_i \mathrm{mES}(\alpha) = \\ -\mu_i - \frac{\partial_i m_2}{2\sqrt{m_2}} E_{G_2}[z|z \leq g_\alpha] + \sqrt{m_2}\frac{1}{\alpha}\{\frac{1}{24}[I^4 - 6I^2 + 3\phi(g_\alpha)]\partial_i k_p \\ + \frac{1}{6}[I^3 - 3I^1]\partial_i s_p + \frac{1}{36}[I^6 - 15I^4 + 45I^2 - 15\phi(g_\alpha)]s_p \partial_i s_p \\ + \partial_i g_\alpha[-g_\alpha\phi(g_\alpha) + \frac{1}{24}[\partial_i I^4 - 6\partial_i I^2 - 3g_\alpha\phi(g_\alpha)]k_p + \frac{1}{6}[\partial_i I^3 - 3\partial_i I^1]s_p \\ + \frac{1}{72}[\partial_i I^6 - 15\partial_i I^4 + 45\partial_i I^2 + 15g_\alpha\phi(g_\alpha)]s_p^2]\}. \end{aligned} \tag{3.2.13a}$$

If we let $z_\alpha = \Phi^{-1}(\alpha)$, then

$$\partial_i g_\alpha = \frac{1}{6}(z_\alpha^2 - 1)\partial_i s_p + \frac{1}{24}(z_\alpha^3 - 3z_\alpha)\partial_i k_p - \frac{1}{18}(2z_\alpha^3 - 5z)s_p\partial_i s_p. \tag{3.2.13b}$$

For q even, we get

$$\partial_i I^q = \sum_{i=1}^{q/2}\left(\frac{\prod_{j=1}^{q/2} 2j}{\prod_{j=1}^{i} 2j}\right) g_\alpha^{2i-1}(2i - g_\alpha^2)\phi(g_\alpha) - \left(\prod_{j=1}^{q/2} 2j\right) g_\alpha\phi(g_\alpha) \tag{3.2.13c}$$

and for q odd

$\partial_i I^q =$

$$\sum_{i=0}^{q^*}\left(\frac{\prod_{j=0}^{q^*}(2j+1)}{\prod_{j=0}^{i}(2j+1)}\right)g_\alpha^{2i}(2i+1-g_\alpha^2)\phi(g_\alpha)-\left(\prod_{j=0}^{q^*}(2j+1)\right)\phi(g_\alpha), \qquad (3.2.13d)$$

with $q^* = (q-1)/2$. (see Boudt et al. 2008, 100.) These resulting formulae are then translated into an efficient algorithm[6] included in the *PortfolioAnalytics* package (see Peterson et al. 2018), and a constraint to minimize component contribution (i.e., equal risk contribution) is specified.[7]

Before solving these four portfolio optimization problems, we add the commonly used long-only and full investment constraints, meaning that short selling is not allowed ($\omega_i \geq 0,\ \forall i \in [1; N]$), and the maximum sum of weights must not exceed one ($\boldsymbol{\omega}'\mathbf{1} = 1$). The backtests are run such that the portfolios are rebalanced monthly without an initial training period or rolling window. Furthermore, in order to produce more realistic performance metrics, we add a proportional transaction cost constraint to each optimization problem and assume a rate of 50 basis points (i.e., 0.5 percent). Our choice is in line with Mei and Nogales (2018), who use the same value and note that this type of transaction cost is realistic to model small trades, where the transaction costs come from the bid-ask spreads, poor liquidity, and various brokerage commissions and fees. Furthermore, since the bid-ask spread is commonly used as a proxy for transaction costs (see, e.g., Caporale & Zakirova 2017, 103), our assumption should be conservative enough, given that the average bid-ask spread of the 40 constituent stocks of the MICEX Index was only about 0.26 percent during the time period considered.

To illustrate how the proportional transaction cost constraint operates, let us assume that transaction costs are paid at the beginning of the period and that the costs associated with transactions must be paid out of the existing budget. In other words, they must be financed from asset sales. Hence, we obtain the following formula:

$$\sum_{i=1}^{n}(\omega_i^- - \omega_i^+) - \sum_{i=1}^{n}(tc_i^+\omega_i^+ + tc_i^-\omega_i^-) \geq 0, \qquad \omega_i^+ \geq 0, \qquad \omega_i^- \geq 0. \qquad (3.2.14)$$

In Equation (3.2.14), the positive weight changes ω_i^+ are assets bought, and the negative weight changes ω_i^- are assets sold. Moreover, $tc_i^{\pm}$ are the proportional transaction costs for buying and selling. The first summation denotes the proceeds from net selling, while the second summation denotes the associated costs. As noted by Scherer and Martin (2005, 100–102), transaction costs lead to an indirect return reduction because the amount on which asset returns can be earned is reduced from the start of the investment period. This can be denoted by $\sum_i \omega_i < \sum_i \omega_i^{initial} = 1$, where $\omega_i^{initial}$ are the equally weighted initial holdings. We can incorporate Equation (3.2.14) into a budget constraint by adding transaction costs to the summation of holdings that are left after transaction costs have

[6] This algorithm solves minimum ES problems via mixed-integer linear programming and can thus compute mES and component mES even for portfolios with a vast number of assets (see Boudt, Carl & Peterson 2010).
[7] Alternatively, one could specify upper and lower bounds on percentage risk contribution.

been paid. The resulting minimization problem, where $\boldsymbol{\mu}$ and σ_{ij} reflect portfolio return and the covariance between assets, is given as follows:

$$\min \sum_i \sum_j \omega_i \omega_j \sigma_{ij} \text{ subject to} \tag{3.2.15a}$$

$$\sum_i \omega_i (1 + \mu_i) = 1 + \boldsymbol{\mu} \tag{3.2.15b}$$

$$\sum_{i=1}^{n} \omega_i + \sum_{i=1}^{n} (tc_i^+ \omega_i^+ + tc_i^- \omega_i^-) = 1 \tag{3.2.15c}$$

$$\omega_i = \omega_i^{initial} + \omega_i^+ - \omega_i^- \tag{3.2.15d}$$

$$\omega_i \geq 0 \tag{3.2.15e}$$

$$\omega_i^+ \geq 0 \tag{3.2.15f}$$

$$\omega_i^- \geq 0. \tag{3.2.15g}$$

For simplicity, we assume that transaction costs are equal across assets.[8] It is also worth mentioning that as the value of proportional transaction costs increases, the constraint pushes the optimal portfolio to the initial weights in order to limit these costs.[9]

The final thing we need to consider before running the backtest is the solver. Here, we utilize a stochastic global optimization algorithm termed Differential Evolution (DE), which is included in the *DEoptim* package (see Ardia et al. 2020). DE is a high-performance optimizer that has shown superior performance in various real-world applications in terms of solution accuracy and repeatability (see, e.g., Vesterstrom & Thomsen 2004; Dong 2009; Iwan et al. 2012; Zhang & Wei 2014). However, the utilization of *DEoptim* becomes more and more time-consuming and computationally burdensome as the portfolio size increases. Thus, for a portfolio consisting of, say, hundreds of assets, one might wish to use, for example, particle swarm optimization instead (see Bendtsen 2015).[10]

Let us next turn to robust covariance estimators and robust portfolio optimization. Recall that, in the above allocation strategies, we followed the typical practice where the arithmetic mean and the sample covariance were used to estimate the population's theoretical location or dispersion moments. Although these classical estimators are asymptotically efficient and consistent, they lose their attractive properties if the distributional assumption is violated. Indeed, due to estimation error, the use of sample estimators for the expected returns and the covariance matrix tends to lead to sub-optimal portfolio solutions. According to Pfaff (2016, 163), sudden swings in the asset mix and extreme weights are also typical in ex-post simulations.

Statistically, these issues are caused mainly by the sample estimators' sensitivity to extreme observations (or outliers). Since the outlying data points tend to have a smaller effect on the dispersion estimates than to the means, *ceteris paribus*, minimum-variance

---

[8] One could change this by providing vectors of transaction costs instead, together with the relevant indexing (see Scherer & Martin 2005, 100–102 for more technical details).

[9] From a technical point of view, a proportional transaction cost constraint can be considered as a penalty for global numeric solvers (see Bennett 2015a, 7–8; Chen et al. 2019, 2–3).

[10] See also Bennett (2015a, 20–25) for more information on this heuristic optimization algorithm, as well as other solvers written in R.

portfolios are often preferred to mean-risk portfolios (see, e.g., Kempf & Memmel 2006). Overall, one would wish to use estimators that reduce the influence of outliers and thus generate estimates that represent the majority of sample data. Alternatively, it would be ideal to use such optimization methods that directly include estimation errors. As noted by Pfaff (2016, 163–164), one can achieve the former by employing robust statistics and the latter by utilizing robust optimization techniques.

To investigate these two approaches, we construct applications in the form of backtest comparisons where the robust optimizations are compared to portfolio solutions based on ordinary sample estimators. In our first application, we optimize minimum-variance portfolios using the classical estimators and the *M*-, *MM*, *Minimum Volume Ellipsoid* (MVE), *Minimum Covariance Determinant* (MCD), *S*-, and *Orthogonalized Gnanadesikan-Kettenring* (OGK) robust estimators for the unknown parameters of location and scatter. We introduce these robust estimators briefly below. The computations are done in R using the *rrcov* (see Todorov 2020) and *cccp* (see Pfaff et al. 2020) packages. The former provides the estimators, whereas the latter is employed to carry out the optimization.

In order to determine the weights of the minimum-variance portfolio for a given estimator, we follow (Pfaff 2016, 181–182, 186–189) and create an estimating function that returns the dispersion estimates (i.e., the second moments) for the classical and robust methods. We apply this function to the empirical data set and use it for portfolio optimization. We can then calculate the portfolio weights for each of the estimation methods by utilizing a rolling window approach of 252 observations. Next, we compute the portfolio returns (for each of the underlying estimators) such that the weights (which are lagged by one period) are multiplied by the respective returns. Finally, to analyze relative performance, we compute the portfolios' excess returns based on robust estimators relative to the classical covariance estimator.

The following brief depiction of the M- and MM estimators is restricted to univariate samples for the sake of convenience, and the notation is the same as used in Pfaff (2016, 165–166).[11] Firstly, the M estimator, which closely resembles the maximum likelihood (ML) principle[12] (see Huber 1964; 1981), is defined as the minimum of a sum of functions $\rho(x, \theta)$:

$$\hat{\theta} = \arg\min\left(\sum_{i=1}^{n} \rho(x_i, \theta)\right). \tag{3.2.16}$$

[11] For more information on the M-estimators, see, e.g., Lopuhaä (1989), Rocke (1996), Todorov and Filzmoser (2009), Maronna et al. (2019, 30, 40–42, 202–204), and Todorov (2020, 17–20, 38–40).

[12] As stated by the ML principle, the unknown parameters (θ) are defined such that they have most likely generated a given i.i.d. sample, $\{x_1, \ldots, x_n\}$, drawn from a distribution $F(x, \theta)$ with density function $f(\cdot)$. Because of this i.i.d. assumption, the joint distribution, which is maximized, is equal to the product of the marginals. However, since the logarithm is a strictly monotone transformation, the optimization can also be executed by minimizing the negative log-likelihood function. (see Pfaff 2016, 165.)

The function $\rho(\cdot)$ has to fulfill the requirements of symmetry, positive definiteness, and a global minimum of zero. Although different specifications of $\rho(\cdot)$ exist,[13] we employ the bi-square (aka bi-weight) function proposed by Tukey (1962):

$$\rho_k(x) = \begin{cases} 1 - [1 - (x/k)^2]^3 & \text{if } |x| \leq k, \\ 1 & \text{if } |x| > k. \end{cases} \quad (3.2.17)$$

As can be seen, this function is bounded for large absolute values of x. Hence, extreme data points receive a smaller weight and are thus less influential regarding the parameter estimates. Next, the algorithm is set to seek a solution to Equation (3.2.16) by equating the gradient $\psi(x) = \delta\rho(x,\theta)/\delta\theta$ to zero. Finally, the MM estimator[14] can be computed by iterating the above method. Indeed, an M-estimator is first used to estimate the scatter of the data. Consistent with the variance-covariance matrix of the first step, a second M-type estimation is then conducted for the location parameters. (Pfaff 2016, 166.) Therefore, the MM-estimator is similar to the generalized M-estimators, i.e., it has the highest possible breakdown point[15] (BP), 0.5, and high efficiency[16] under normality. On the other hand, it can be sensitive to the contamination bias. Moreover, situations, where the iterative estimation procedure does not converge are common. (Wilcox 2012, 499; see also Tatsuoka & Tyler 2000; Salibián-Barrera, Van Aelst & Willems 2006; Maronna et al. 2019, 128–133 for further discussions.)

The MVE, MCD, and S-estimators are obtained by robust scaling. Here, the starting point is a p-dimensional random variable $\boldsymbol{x} = (x_1, \dots, x_p)' \in \mathbb{R}^p$ that is assumed to be jointly normally distributed, i.e., $\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$. Thus, the location vector is $\boldsymbol{\mu} = \mathrm{E}(\boldsymbol{x}) = \left(\mathrm{E}(x_1), \dots, \mathrm{E}(x_p)\right)'$ and the covariance matrix is $\Sigma = \mathrm{VAR}(\boldsymbol{x}) = \mathrm{E}\left((\boldsymbol{x} - \boldsymbol{\mu})(\boldsymbol{x} - \boldsymbol{\mu})'\right)$, where VAR denotes the variance. Moreover, the squared Mahalanobis distance between the sample and the location and dispersion parameters is used as the distance measure such that $d(\boldsymbol{x}, \boldsymbol{\mu}, \Sigma) = (\boldsymbol{x} - \boldsymbol{\mu})'\Sigma^{-1}(\boldsymbol{x} - \boldsymbol{\mu})$. In order to exclude the trivial solution of zero distance, we require that $|\hat{\Sigma}| = 1$, where $\hat{\Sigma}$ is the estimated covariance matrix and $|\cdot|$ is a matrix determinant. Each of these three robust estimators for $\hat{\boldsymbol{\mu}}$ and $\hat{\Sigma}$ is then based on the following optimization:

---

[13] Alternatively, one could use the class of Huber functions. However, the bi-square function is a better option for symmetric distributions characterized by excess kurtosis since the impact of outliers can be completely suppressed. (Pfaff 2016, 166; see also Todorov 2020 for more details.)

[14] For more discussions on the class of MM estimators, see, e.g., Yohai (1987), Yohai, Stahel, and Zamar (1991), as well as Lopuhaä (1991; 1992).

[15] The breakdown point is a measure used for estimating the robustness of an estimator. It can take values between 0 and 0.5 and is defined as the relative share of outliers in a sample such that the estimator does not produce an arbitrarily large value (Würtz et al. 2009, 293). In other words, the BP of an estimator $\hat{\theta}$ of the parameter θ is the largest amount of contamination (proportion of atypical points) that the data can contain such that $\hat{\theta}$ still provides information about θ (about the distribution of the 'typical points') (Maronna et al. 2019, 58–59; see also Todorov 2020).

[16] Relative efficiency is used to evaluate the suitability of a robust estimator as it shows how much the sample size has to be increased in order for the variances of the two estimators to be equalized. More exactly, the asymptotic variance of a robust estimator is given relative to the variance of an optimal estimator that has been derived under strict adhesion to the distribution or model assumption. (Pfaff 2016, 165.)

$$\arg\min \hat{\sigma}\left(\boldsymbol{d}\big(X, \hat{\boldsymbol{\mu}}, \hat{\Sigma}\big)\right), \tag{3.2.18}$$

where $\hat{\sigma}$ is a robust metric, $\boldsymbol{d}$ is the vector of distances $d\big(x_i, \hat{\boldsymbol{\mu}}, \hat{\Sigma}\big)$ for $i = p + 1, \ldots, N$ and X is the full sample. (Pfaff 2016, 166–167.)

When utilizing the MVE[17] estimator, samples from a multivariate normal distribution construct ellipsoid-shaped 'clouds' of data points. The smallest data cloud is defined such that at least half of the sample observations are included (i.e., the uncontaminated portion of the data). This is done by using the median of $\hat{\sigma}$. These 'clean' observations, which act as preliminary estimates of the mean vector and the covariance matrix, are then employed to compute a robust Mahalanobis distance for every observation in the sample. Observations for which the robust Mahalanobis distances exceed the 97.5 percent significance level for the χ^2-distribution are marked as probable outliers. (Würtz et al. 2009, 281–284; Pfaff 2016, 167; Todorov 2020, 51.) Unfortunately, the convergence rate of the MVE is only $N^{-1/3}$, and hence it is rather inefficient (see Davies 1992).

With respect to the MCD estimator, a robust subsample of size $h > N/2$ is selected such that the determinant of the variance-covariance matrix is minimal. The location vector is defined as the average (p arithmetic means) of the selected h data points, whereas the estimator for the variance-covariance matrix is multiplied by a consistency factor and a finite sample correction factor. As a result of these adjustments, $\hat{\Sigma}$ follows the normal distribution and is unbiased for small sample sizes. (see Todorov & Filzmoser 2009, 11–13; Würtz et al. 2009, 279; Pfaff 2016, 167; Maronna et al. 2019, 208, 210 for more details.) To compute the MCD estimator, we utilize the fast algorithm developed by Rousseeuw and Van Driessen (1999).[18] Finally, for both the MVE and MCD estimators, we use the default setting for the α parameter, which controls the size of the subsets over which the volume of the ellipsoid and determinant, respectively, is minimized (see Todorov 2020, 37, 52).[19]

It is worth noting that the MVE and the MCD are very drastic because they are intended to safeguard against up to 50 percent of outliers (Rousseeuw & Leroy 1987, 263). It has also been shown that the high breakdown value and low *max bias* (i.e., the maximal possible asymptotic bias) of the MVE estimator make it suitable for outlier detection (Van Aelst & Rousseeuw 2009, 79). On the other hand, even the MCD's asymptotic efficiency is very low (see Paindaveine & Van Bever 2014). To overcome the inefficiencies of these estimators, a more general class of S- (or 'scale') estimators was initially introduced by Davies (1987). In order to perform optimization, an M-estimator for the scale N is inserted to Equation (3.2.18) such that $1/N \sum_{i=1}^{N} \rho(d_i) = \delta$. It is further required that $\delta \in (0,1)$

---

17 See Rousseeuw (1984; 1985) and Rousseeuw and Leroy (1987, 258–265) for more discussions on the MVE and MCD estimators.

18 The underlying principles of this algorithm are an inequality involving order statistics and determinants, as well as techniques called 'selective iteration' and 'nested extensions' (see Hubert, Rousseeuw & van Aelst 2008).

19 It relies on the fact that the breakdown point is maximal for $h = \mathrm{int}[(N + p + 1)/2]$ (Pfaff 2016, 167).

and that $\rho(\cdot)$ is a bounded function. (Pfaff 2016, 167; see also Todorov & Filzmoser 2009, 15–16; Maronna et al. 2019, 208–209.) Hence, S-estimates of location and scatter based on the bi-weight function can be computed using a fast algorithm similar to the one proposed by Salibian-Barrera and Yohai (2006) for the case of regression (see Todorov 2020, 34–37, 50–52, 62–65 for more technical details about the algorithms used).

Finally, we use the OGK estimator proposed by Maronna and Zamar (2002) to avoid the issue of non-convex optimization caused by the robust estimators possessing the affine invariance property. In line with the scatter estimation developed by Gnanadesikan and Kettenring (1972), OGK estimates the covariances between two random variables as:

$$s_{jk} = \frac{1}{4}\left(\sigma\left[\frac{Y_j}{\sigma(Y_j)} + \frac{Y_k}{\sigma(Y_k)}\right]^2 - \sigma\left[\frac{Y_j}{\sigma(Y_j)} - \frac{Y_k}{\sigma(Y_k)}\right]^2\right). \tag{3.2.19}$$

Thus, the affine invariance property is replaced by a more straightforward calculation method. The name orthogonalized Gnanadesikan–Kettering (OGK) comes from the fact that an orthogonalization to the data pairs of X must be implemented so that the resulting variance-covariance matrix remains positive definitive. It can be shown that if a robust estimator for the standard deviation σ is utilized for the pairwise covariances s_{jk}, with $j = 1, \dots, p$ and $k = 1, \dots, p$, this matrix is also robust. (Pfaff 2016, 167–168; Maronna et al. 2019, 224–228; see also Todorov & Filzmoser 2009, 14–15; Wilcox 2012, 232–233; Todorov 2020, 53–55 for more details on the computation of the OGK estimator.) According to Maronna and Zamar (2002, 316), OGK performs similarly to Rousseeuw and Van Driessen's (1999) fast MCD estimates with high-dimensional real data.

Let us now turn to our second application, where we construct optimal portfolios using robust techniques. As a motivation, recall that so far, all the resulting estimates have been fixed, and we have utilized them directly as inputs. In reality, however, parameter values are subject to uncertainty. Therefore, it is worth considering an approach that can mitigate the adverse effects caused by different parameter constellations or, more specifically, worst-than-expected returns. This limitation is made because, as is well-known, expected returns tend to have a stronger impact on the portfolio composition than their dispersion.[20]

Following the notation used by Pfaff (2016, 168–169), we first define the classical mean-variance (MV) portfolio optimization by

$$P_\lambda = \underset{\boldsymbol{\omega}\in\Omega}{\arg\min}(1-\lambda)\sqrt{\boldsymbol{\omega}'\Sigma\boldsymbol{\omega}} - \lambda\boldsymbol{\omega}'\boldsymbol{\mu}. \tag{3.2.20}$$

Here, $\boldsymbol{\omega}$ is the $n \times 1$ portfolio weight vector and $\Omega \subset \{\boldsymbol{\omega} \in \mathbb{R}^N | \boldsymbol{\omega}'\mathbf{1} = 1\}$ is the set of all permissible solutions. The (expected) returns of the N stocks are included in the vector

[20] The reader interested in learning more about robust (portfolio) optimization in the presence of uncertainty is referred, e.g., to Lobo and Boyd (2000), Scherer (2002), Beyer and Sendhoff (2007), Michaud and Michaud (2008), Schöttle and Werner (2009), Bertsimas, Brown, and Caramanis (2011), Ye, Parpas and Rustem (2012), Gorissen, Yanıkoğlu and den Hertog (2015), Buhmann et al. (2018), as well as to Yin, Perchet and Soupé (2019).

$\boldsymbol{\mu}$, with variance-covariance matrix $\Sigma \in \mathbb{R}^{N\times N}$.[21] The parameter λ determines the weighting between the portfolio return and its risk, and it can take all values in the interval $[0, 1]$. Of course, if one sets $\lambda = 0$ or $\lambda = 1$, this formulation incorporates the specific cases of a minimum-variance and a maximum-return portfolio, respectively.[22]

Next, we need to incorporate the uncertain return parameters directly in the optimization formulation. Within the robust optimization framework, it is common to define an uncertainty set $\mathcal{U}_{\hat{\boldsymbol{\mu}}}$ for all parameter values allowed. Although there are different types of sets, we settle for using the *elliptical uncertainty set*:

$$\mathcal{U}_{\hat{\boldsymbol{\mu}}} = \left\{\boldsymbol{\mu} \in \mathbb{R}^N \,|\, (\mu - \hat{\mu})'\hat{\Sigma}^{-1}(\mu - \hat{\mu}) \leq \frac{\delta^2}{T}\right\}, \tag{3.2.21}$$

where $\boldsymbol{\mu}$ denotes the expected return vector (i.e., the location parameters) and T denotes the sample size. For this uncertainty set, it is assumed that the uncertainties emerge from a multivariate elliptical distribution. In other words, the covariances between stock returns are explicitly included in the uncertainty set. (see Schöttle 2007, 148–149; Pfaff 2016, 169–170 for more information on alternative specifications.)

In order to express this uncertainty set in a mathematical program, we utilize the *worst-case* (aka *min-max*) approach. It determines the optimal weight vector given the least favorable parameter constellation (i.e., the lowest portfolio return for a given risk level). The resulting optimization problem can now be written as:

$$PR_\lambda = \underset{\boldsymbol{\omega}\in\Omega}{\arg\min}\ \underset{\boldsymbol{\mu}\in\mathcal{U}}{\arg\max}\ (1-\lambda)\sqrt{\boldsymbol{\omega}'\Sigma\boldsymbol{\omega}} - \lambda(\boldsymbol{\omega}'\boldsymbol{\mu}). \tag{3.2.22}$$

It is clear that $(\boldsymbol{\mu} - \hat{\boldsymbol{\mu}})'\Sigma^{-1}(\boldsymbol{\mu} - \hat{\boldsymbol{\mu}})$ is distributed as χ^2 with N degrees of freedom if the returns are multivariate normally distributed. In such a case, the scalar δ^2 is the equivalent quantile value for a given confidence level $(1-\alpha)$. Hence, the stochastic return vector $\boldsymbol{\mu}$ is located in the ellipse determined by the level of confidence, and the maximal distance between the empirical location vector and this uncertainty ellipsoid is defined such that the returns are consistent with the least favorable outcome. (Pfaff 2016, 170–171; Schöttle 2007, 121–128.)

More precisely, as the variance-covariance matrix is taken as given and the uncertainty is limited to the returns only, we can define the maximal distance by using a Lagrange approach where $\hat{P} = \frac{1}{T}\hat{\Sigma}$ is the variance-covariance matrix of the empirical returns, and γ denotes the Lagrangian multiplier:

$$\mathcal{L}(\boldsymbol{\mu}, \gamma) = \boldsymbol{\omega}'\hat{\boldsymbol{\mu}} - \boldsymbol{\omega}'\boldsymbol{\mu} - \left[\frac{\gamma}{2}(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu})'\hat{P}^{-1}(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}) - \delta^2\right]. \tag{3.2.23}$$

In order to find the optimal solution, we take the partial derivatives of (3.2.23) with respect to $\boldsymbol{\mu}$ and γ and set these to zero. This leads to a system of two equations, which we solve for $\boldsymbol{\mu}$:

[21] Further constraints can be added to the formulation in Equation (3.2.20), such as a budget ($\boldsymbol{\omega}'\mathbf{1} = 1$) and non-negativity ($\boldsymbol{\omega} \geq 0$) constraints.

[22] λ values that are between these bounds generate portfolios located on the feasible efficient frontier.

$$\boldsymbol{\mu} = \hat{\boldsymbol{\mu}} - \left(\frac{\delta}{\sqrt{\boldsymbol{\omega}'\hat{P}\boldsymbol{\omega}}} P\boldsymbol{\omega}\right). \tag{3.2.24}$$

Once the left-multiplication by $\boldsymbol{\omega}'$ has been performed, we get the following result for the portfolio returns:

$$\begin{aligned} \boldsymbol{\omega}'\boldsymbol{\mu} &= \boldsymbol{\omega}'\hat{\boldsymbol{\mu}} - \delta\sqrt{\boldsymbol{\omega}'\hat{P}\boldsymbol{\omega}} \\ &= \boldsymbol{\omega}'\hat{\boldsymbol{\mu}} - \delta\left\|\hat{P}^{\frac{1}{2}}\boldsymbol{\omega}\right\|. \end{aligned} \tag{3.2.25}$$

As can be seen, the portfolio return is *smaller* than the classical solution by the term $\delta\sqrt{\boldsymbol{\omega}'P\boldsymbol{\omega}}$ when robust optimization with elliptical uncertainty is utilized. Moreover, the square root of the quantile value can be understood as a risk aversion parameter in regard to uncertain estimates. Finally, if we substitute the inner solution from Equation (3.2.25) into a robust optimization formulation similar to the one in Equation (3.2.22), we get:

$$\begin{aligned} PR_\lambda &= \underset{\boldsymbol{\omega}\in\Omega}{\arg\min}\ \underset{\boldsymbol{\mu}\in\mathcal{U}}{\arg\max}\ (1-\lambda)\sqrt{\boldsymbol{\omega}'\Sigma\boldsymbol{\omega}} - \lambda(\boldsymbol{\omega}'\boldsymbol{\mu}) \\ &= \underset{\boldsymbol{\omega}\in\Omega}{\arg\min}\ (1-\lambda)\sqrt{\boldsymbol{\omega}'\Sigma\boldsymbol{\omega}} - \lambda(\boldsymbol{\omega}'\boldsymbol{\mu}) + \lambda\frac{\delta}{\sqrt{T}}\sqrt{\boldsymbol{\omega}'\hat{\Sigma}\boldsymbol{\omega}} \\ &= \underset{\boldsymbol{\omega}\in\Omega}{\arg\min}\ \left(1-\lambda+\lambda\frac{\delta}{\sqrt{T}}\right)\sqrt{\boldsymbol{\omega}'\Sigma\boldsymbol{\omega}} - \lambda\boldsymbol{\omega}'\hat{\boldsymbol{\mu}}. \end{aligned} \tag{3.2.26}$$

This equation has a couple of important implications. Firstly, if robust optimization under elliptical uncertainty is employed, a portfolio's *efficient frontier* is the same as the efficient frontier of a classical MV portfolio, excluding a shortening factor. Secondly, as the uncertainty is limited to returns only, the optimal weight vector for a *minimum-variance portfolio* is the same for optimization methods. (see Schöttle 2007, 124–126; Pfaff 2016, 172.)

Let us focus on the first implication and express the risk-return trade-off parameter in (3.2.26) as θ. Hence, the corresponding trade-off parameter λ in the problem specification of (3.2.20) is given by

$$\lambda := \frac{\theta}{1+\theta\frac{\delta}{\sqrt{T}}}. \tag{3.2.27}$$

We can now see that the determined interval for $\theta \in [0,1]$ is bounded above for the equivalent classical MV portfolios as $\lambda \in \left[0, 1/\left(1+\delta/\sqrt{T}\right)\right]$. Overall, the solution in the case of an elliptical uncertainty set is conservative by design because the optimal portfolio weights are consistent with a situation where the least favorable return is realized for every stock. (Pfaff 2016, 172–173.)

Next, we need to express the problem formulation in (3.2.22) as a Second-Order Cone Program (SOCP).[23] Thus, we introduce a slack variable $t = \boldsymbol{\mu}'\boldsymbol{\omega}$ for the unknown worst-case portfolio return and write the problem as follows:

[23] SOCPs comprise linear and quadratic programs, problems with hyperbolic constraints, as well as problems involving sums or maxima of vector norms. They can be solved by utilizing interior-point methods. (Pfaff 2016, 173.) For thorough theoretical expositions, see, e.g., Nesterov and Nemirovskii (1994), Lobo et al. (1998), Lobo (2000), Boyd and Vandenberghe (2004).

$$\underset{t,\boldsymbol{\omega}}{\arg\min}\ t \tag{3.2.28a}$$

$$\text{subject to} \quad t \leq \hat{\boldsymbol{\mu}}'\boldsymbol{\omega} - \delta_\alpha \left\| P^{\frac{1}{2}}\boldsymbol{\omega} \right\|, \tag{3.2.28b}$$

$$\sigma_{\max} \geq \left\| \Sigma^{\frac{1}{2}}\boldsymbol{\omega} \right\|. \tag{3.2.28c}$$

The first inequality constraint is the cone constraint, and the second is a quadratic constraint regarding the portfolio risk.[24] Naturally, the above SOCP form can be extended by other linear constraints (e.g., non-negativity, budget, group, and bound constraints). (Pfaff 2016, 173–174.)

In Subchapter 4.1, we compare the efficient frontiers for a classical MV optimization as given in (3.2.20) and the robust counterpart (RC) optimization as given in (3.2.26) when the elliptical uncertainty is based on the maximum likelihood point estimates.[25] Both problems are used to maximize the expected portfolio return for a given portfolio standard deviation risk. We also add the long-only (non-negativity) and full investment (budget) constraints. To solve the SOCP, we utilize the *cccp* ('cone constrained convex programs') package (see Pfaff et al. 2020). Consistent with Pfaff (2016, 190–194), we set the desired confidence level to 90 percent and determine the robust allocation for a value of $\theta = 0.7$.[26] Next, we define a sequence of ten risk aversion parameters for computing the points on the efficient frontier. These are created such that the range covers the portfolio standard deviation risks greater than 110 percent of the least risky and 90 percent of the riskiest stock in terms of the standard deviation risk. Finally, we provide a graphical representation of the efficient frontier with equivalence points.

The above robust optimization example can be developed further. Indeed, instead of only investigating the portfolio allocations along the efficient frontier for a fixed point in time, we also wish to study the cumulated out-of-sample behavior of robustly optimized portfolios and compare their performance to the selected benchmark portfolios. Therefore, we utilize MATLAB® R2021a and its Financial Toolbox™ to construct five different investment strategies and execute the backtests over a roughly five-year period.

Our first simple benchmark strategy invests equally across all stocks. Following MATLAB's mathematical notation (see Backtest investment strategies 2021), this *equal-weighted* strategy can be formulated as:

$$\boldsymbol{\omega}\text{EW} = (\omega_1, \omega_2, \dots, \omega_N), \qquad \omega_i = \frac{1}{N}. \tag{3.2.29}$$

---

[24] In the case of a quadratic cone optimization, a required property for a corresponding cone (aka an ice-cream or Lorenz cone) is that the first element is at least as great as the Euclidean norm of its remaining elements. Mathematically, the cone can be defined as $\mathcal{C} = \{\boldsymbol{x} = (x_1, \dots, x_N) \in \mathbb{R}^N : x_1 \geq \|x_2, \dots, x_N\|\}$, where $\boldsymbol{x}$ is a vector. In our case, $\boldsymbol{x} = (t, \boldsymbol{\omega})'$. (see Pfaff 2016, 173.)

[25] Recall that these two problem formulations only differ with respect to the factor in the first term. Pfaff (2016, 190–194) calls this the 'sigma term'.

[26] Note that we do not need to provide the risk-weighting parameter λ since it is a scaler that does not affect the optimization outcome but only the value of the objective function (Pfaff 2016, 191).

The second strategy is the *maximization of the Sharpe ratio*, where the objective is to maximize mean return per unit standard deviation as follows:

$$\boldsymbol{\omega}\mathrm{SR} = \arg\max_{\boldsymbol{\omega}} \left\{ \frac{\boldsymbol{r}'\boldsymbol{\omega}}{\sqrt{\boldsymbol{\omega}'Q\boldsymbol{\omega}}} \mid \boldsymbol{\omega} \geq 0, \sum\nolimits_{1}^{N} \omega_i = 1,\ 0 \leq \boldsymbol{\omega} \leq 0.25 \right\}. \qquad (3.2.30)$$

Here, $\boldsymbol{r}$ is a vector of expected returns, and Q is the covariance matrix of asset returns. The long-only and full investment constraints are also imposed. The upper bound of 0.25 means that a maximum of twenty-five percent is invested into a single asset (including cash). The maximum Sharpe ratio portfolio (MSRP) can be considered as a portfolio somewhere between the minimum variance portfolio (MVP) and maximum return portfolio (MRP). Hence, one should expect the effects to be similar to those occurring in the case of the MRP but milder (see Schöttle 2007, 154–155).

The third competing asset allocation approach is the *inverse variance* strategy:

$$\boldsymbol{\omega}\mathrm{IV} = (\omega_1, \omega_2, \ldots, \omega_N), \qquad \omega_i = \frac{(\sigma_{\mathrm{ii}}^{-1})}{\sum_{i=1}^{N} \sigma_{\mathrm{ii}}^{-1}}, \qquad (3.2.31)$$

where σ_{ii} are diagonal elements of the asset return covariance matrix. In this case, the risk is measured with variance, and assets are weighted in inverse proportion to their risk (see, e.g., Maillard et al. 2010; Arévalo, León & Hernandez 2019 for further discussion). The fourth strategy considers the *Markowitz portfolio optimization*, i.e., maximizing return and minimizing risk with a fixed risk aversion coefficient:

$$R_{\mathrm{Mkwtz}} = \max_{\boldsymbol{\omega}} \left\{ \boldsymbol{r}'\boldsymbol{\omega} - \lambda\boldsymbol{\omega}'Q\boldsymbol{\omega} \mid \boldsymbol{\omega} \geq 0, \sum\nolimits_{1}^{N} \omega_i = 1, 0 \leq \boldsymbol{\omega} \leq 0.25 \right\}, \qquad (3.2.32)$$

where $\lambda = 0.05$ is the risk aversion coefficient. Of course, this is the classical quadratic utility maximization problem, where the λ parameter is used to control how much portfolio variance is penalized.[27]

The fifth strategy assigns asset weights using *robust optimization* with uncertainty in expected returns.[28] Thus, instead of modeling unknowns as one point, we specify them as a set of values containing the most likely possible realizations. More precisely, $\boldsymbol{r} = \{\boldsymbol{r} | \boldsymbol{r} \in S(\boldsymbol{r}_0)\}$, where the expected return is defined not by the deterministic vector $\boldsymbol{r}_0$ but by the region $S(\boldsymbol{r}_0)$ around the vector $\boldsymbol{r}_0$. The portfolio selection is now formulated as a problem of finding the maximum and minimum as follows:

$$R_{\mathrm{robust}} = \max_{\boldsymbol{\omega}} \min_{r \in S(r_0)} \left\{ \boldsymbol{r}'\boldsymbol{\omega} - \lambda\boldsymbol{\omega}'Q\boldsymbol{\omega} \mid \boldsymbol{\omega} \geq 0, \sum\nolimits_{1}^{N} \omega_i = 1, 0 \leq \boldsymbol{\omega} \leq 0.25 \right\}. \qquad (3.2.33)$$

The region of uncertainty $S(\boldsymbol{r}_0)$ is again specified as an ellipsoid so that

$$S(\boldsymbol{r}_0) = \{\boldsymbol{r} | (\boldsymbol{r} - \boldsymbol{r}_0)'\Sigma_r^{-1}(\boldsymbol{r} - \boldsymbol{r}_0) \leq \kappa^2\}. \qquad (3.2.34)$$

Here, $\kappa = 1.1$ is the (robust) uncertainty aversion coefficient defining how wide the uncertainty region is, and Σ_r is the matrix of estimation errors in expected returns $\boldsymbol{r}$. With

[27] See, e.g., Bodnar et al. (2018) for more discussion on the estimation of risk aversion coefficients.

[28] More examples of various robust asset allocation problems can be found, e.g., in Goldfarb and Iyengar (2003), Tütüncü and Koenig (2004), Scherer and Martin (2005, 238–245), as well as in Cornuéjols et al. (2018, 289–302).

the addition of the ellipsoid uncertainty to the Markowitz model, we can reformulate the robust optimization problem as follows:

$$R_{\text{robust}} = \max_{\boldsymbol{\omega}} \{\boldsymbol{r}'\boldsymbol{\omega} - \lambda\boldsymbol{\omega}'Q\boldsymbol{\omega} - \text{kz} \mid \boldsymbol{\omega} \geq 0,\ z \geq 0,\ \boldsymbol{\omega}'\Sigma_r\boldsymbol{\omega} - z^2 \leq 0,\ \textstyle\sum_1^N \omega_i = 1,\ 0 \leq \boldsymbol{\omega} \leq 0.25\}. \tag{3.2.35}$$

(cf. Backtest investment strategies 2021.)

Next, we compute the initial weights for each portfolio.[29] This initialization consists of the first 252 days, and the backtest is then run over the remaining data. The portfolios are rebalanced approximately every month $(252/12 = 21)$, and the size of the rolling lookback window is set to 252 days. We also incorporate the fees that the strategies pay for buying and selling stocks by setting the transaction cost rate to one percent (100 basis points).[30] These proportional costs are then deducted from the gross portfolio returns to calculate the net portfolio returns.[31] Furthermore, we assume that the portfolio's value is 10,000,000 monetary units at the start of the backtest, and, if the sum of portfolio weights is below one, the remaining capital can be invested in cash earning risk-free interest percent per annum.[32] Finally, we create a summary table and graph of the performance results. (see Backtest investment strategies 2021; Portfolio optimization theory 2021 for more technical details.)

In order to analyze the *ex-ante* riskiness of these five strategies, we conduct a simple Value-at-Risk estimation at the 95 percent confidence level using the Exponential Weighted Moving Average (EWMA) method.[33] By definition, this method assigns non-equal, exponentially decreasing weights, meaning that the most recent returns have higher weights because they influence today's return more heavily than returns further in the past. For the sake of convenience, an infinitely large estimation window is assumed to approximate the variance:

$$\hat{\sigma}_t^2 \approx (1-\lambda)\left(y_{t-1}^2 + \sum_{i=2}^{\infty} \lambda^{i-1} y_{t-i}^2\right) = (1-\lambda)y_{t-1}^2 + \lambda\hat{\sigma}_{t-1}^2. \tag{3.2.36}$$

[29] Note that if these initial weights were not set, the strategies would begin the backtest uninvested, i.e., 100 percent of the capital would be in cash, earning the risk-free rate, until the first rebalance date.

[30] In MATLAB®, transaction costs are paid as a percentage of the total change in position for each stock. Thus, in contrast to the optimization constraint used in R, a simple scalar decimal percentage charge is used for both purchases and sales of securities.

[31] Alternatively, one could consider scaled transaction costs based on the change in the market value of each asset after a rebalance. In such a case, costs would be computed at separate buy and sell rates, which, e.g., decrease as the changes in asset positions for all assets (in currency units) increase over certain thresholds.

[32] It would also be possible to let the portfolio weights sum to more than one. Allowing this would be analogous to borrowing capital on margin to invest with leverage (i.e., a negative cash position).

[33] Alternatively, one could use, e.g., the normal distribution (parametric) or the historical simulation (non-parametric) method. However, those assume that all past returns carry the same weight. Moreover, the EWMA generally reacts rapidly to periods of large (or small) returns and follows the trend of returns more closely and more accurately than the normal distribution approach or the historical simulation method. Hence, EWMA tends to have fewer VaR violations. (see Value-at-Risk estimation and backtesting 2021.)

In line with the common practice, we set the value of the decay factor to 0.94. The test window starts on the first day in 2014, running through the end of the sample.[34] (see Value-at-Risk estimation and backtesting 2021 for more technical details.)

Finally, with respect to *ex-post* risk analysis, we compute and report the *historical* and *modified VaR*, as well as the *historical* and *modified Expected Shortfall*. The historical VaR refers to a non-parametric VaR estimation method that uses the historical return distribution and its probability quantile. Thus, the historical VaR at a probability level p (here, 95 percent) is the p-quantile of the negative returns, or equivalently, is the negative value of the $c = 1 - p$ quantile of the returns. Analogously, the historical ES is estimated by the negative value of the sample average of all returns that are below the $c = 1 - p$ empirical quantile. The calculation of the modified VaR, in turn, incorporates the higher moments of non-normal distributions via an analytical estimation using a Cornish-Fisher expansion. In other words,

$$\text{Modified VaR} = -\bar{R} - \sqrt{(\sigma)} \cdot z_{cf}, \tag{3.2.37a}$$

where

$$z_{cf} = z_c + \frac{(z_c^2 - 1)S}{6} + \frac{(z_c^3 - 3z_c)K}{24} - \frac{(2z_c^3 - 5z_c)S^2}{36}. \tag{3.2.37b}$$

Here, $\bar{R}$ is the mean return, σ is the variance of the return distribution R, S is the skewness of R, K is the excess kurtosis of R, and z_c is the c-quantile of R. The calculation of the modified ES follows the method by Boudt et al. (2008) discussed earlier. Note that although modified ES should always be larger than modified VaR, this may not always be the case due to estimation problems. (see Peterson et al. 2015, 11–12, 107–108, 228–230 and references therein for more technical details.)

3.3 Applications of the higher-order comoments for portfolio selection

As it is well established, the classical mean-variance portfolio selection techniques can involve a severe welfare loss in the presence of non-normally distributed returns and non-quadratic preferences (see, e.g., Martellini & Ziemann 2008, 3 and references therein). Therefore, in Subchapter 4.2, we focus on two cases where the objective is to minimize downside risk or maximize expected utility and employ the asset allocation framework proposed by Boudt, Lu, and Peeters (2015). This general framework improves estimates of the variance-covariance matrix by introducing enhanced estimates of the co-skewness and co-kurtosis parameters. Here, explicit formulas for the higher-order comoments are derived, assuming that a multifactor model is an underlying data-generating process of

[34] Due to space limitations, we omit further VaR backtesting analyses and refer the interested reader, e.g., to Abad, Benito, and López (2014) for a comprehensive review of methodologies developed to estimate the VaR. Correspondingly, more information on financial risk forecasting with MATLAB® can be found in Danielsson (2011).

stock returns. Naturally, such an assumption also reduces the number of parameters to estimate compared to the traditional approach.

In line with Boudt et al. (2015, 226–227), we stack all the comoments into a $N \times N$ covariance matrix Σ, $N \times N^2$ coskewness matrix Φ, and $N \times N^3$ cokurtosis matrix Ψ of the return vector $\boldsymbol{r}$ (with mean $\boldsymbol{\mu}_r$) as follows:

$$\Sigma = E[(\boldsymbol{r} - \boldsymbol{\mu}_r)(\boldsymbol{r} - \boldsymbol{\mu}_r)'], \tag{3.3.1a}$$

$$\Phi = E[(\boldsymbol{r} - \boldsymbol{\mu}_r)(\boldsymbol{r} - \boldsymbol{\mu}_r)' \otimes (\boldsymbol{r} - \boldsymbol{\mu}_r)'], \tag{3.3.1b}$$

$$\Psi = E[(\boldsymbol{r} - \boldsymbol{\mu}_r)(\boldsymbol{r} - \boldsymbol{\mu}_r)' \otimes (\boldsymbol{r} - \boldsymbol{\mu}_r)' \otimes (\boldsymbol{r} - \boldsymbol{\mu}_r)'], \tag{3.3.1c}$$

where $\otimes$ denotes the Kronecker product. These comoments enable the calculation of the k-th portfolio return moment for a portfolio with weights $\boldsymbol{\omega}$:

$$m_2(\boldsymbol{\omega}) = E\left[\left(\boldsymbol{\omega}'(\boldsymbol{R} - \boldsymbol{\mu}_R)\right)^2\right] = \boldsymbol{\omega}'\Sigma\boldsymbol{\omega}, \tag{3.3.2a}$$

$$m_3(\boldsymbol{\omega}) = E\left[\left(\boldsymbol{\omega}'(\boldsymbol{R} - \boldsymbol{\mu}_R)\right)^3\right] = \boldsymbol{\omega}'\Phi(\boldsymbol{\omega} \otimes \boldsymbol{\omega}), \tag{3.3.2b}$$

$$m_4(\boldsymbol{\omega}) = E\left[\left(\boldsymbol{\omega}'(\boldsymbol{R} - \boldsymbol{\mu}_R)\right)^4\right] = \boldsymbol{\omega}'\Psi(\boldsymbol{\omega} \otimes \boldsymbol{\omega} \otimes \boldsymbol{\omega}). \tag{3.3.2c}$$

However, due to the curse of dimensionality, the use of unrestricted estimators of the first four (co-)moments is almost infeasible for moderately large dimensions. Moreover, since the optimizer amplifies the estimation errors, one may end up with poorly diversified portfolios (see, e.g., Green & Hollifield 1992). Thus, it is reasonable to impose more structure on the data through a (statistical) factor model, under which the variation in asset returns is assumed to be driven by multiple common factors and idiosyncratic factors specific to each asset.

Indeed, let us assume that K observable factors are identified as being influential for the portfolio variability and that, at a given frequency, the asset returns $\boldsymbol{r}_t = (r_{1t}, \dots, r_{Nt})'$ and the factors $\boldsymbol{f}_t = (f_{1t}, \dots, f_{Kt})'$ are observed. The asset returns are assumed to depend linearly on the factors, whereby the variation in the asset returns that is not explained by the factors (i.e., the residual terms, $\boldsymbol{e}_t$), is assumed to be independent of each of the factors and independent across assets. Such a system can be presented in matrix notation as:

$$\boldsymbol{r}_t = a + B\boldsymbol{f}_t + \boldsymbol{e}_t, \tag{3.3.3}$$

where $\boldsymbol{e}_t = (e_{1t}, \dots, e_{Nt})'$ is the $N \times 1$ vector of independent asset-specific factors, and B is the $N \times K$ matrix of factor loadings (i.e., the factor betas or factor exposures) of the N assets on the K factors. (Boudt et al. 2015, 226–228.)

The comoments under the linear factor model are defined as follows: Let S be the $K \times K$ covariance matrix of the K factors, G the $K \times K^2$ coskewness matrix of the K factors and P their $K \times K^3$ cokurtosis matrix. Therefore,

$$\boldsymbol{\mu}_f = E[\boldsymbol{f}_t], \tag{3.3.4a}$$

$$S = E\left[(\boldsymbol{f}_t - \boldsymbol{\mu}_f)(\boldsymbol{f}_t - \boldsymbol{\mu}_f)'\right], \tag{3.3.4b}$$

$$G = E\left[(\boldsymbol{f}_t - \boldsymbol{\mu}_f)(\boldsymbol{f}_t - \boldsymbol{\mu}_f)' \otimes (\boldsymbol{f}_t - \boldsymbol{\mu}_f)'\right], \tag{3.3.4c}$$

$$P = E\left[(\boldsymbol{f}_t - \boldsymbol{\mu}_f)(\boldsymbol{f}_t - \boldsymbol{\mu}_f)' \otimes (\boldsymbol{f}_t - \boldsymbol{\mu}_f)' \otimes (\boldsymbol{f}_t - \boldsymbol{\mu}_f)'\right]. \tag{3.3.4d}$$

The comoment matrices Σ, Φ, and Ψ, are now rewritten as the sum of the comoment of the return explained by the factor (i.e., the comoment of $B\boldsymbol{f}_t$) and a residual matrix (denoted by Δ, Ω, and ϒ) so that

$$\Sigma = BSB' + \Delta, \tag{3.3.5a}$$

$$\Phi = BG(B \otimes B') + \Omega, \tag{3.3.5b}$$

$$\Psi = BP(B' \otimes B' \otimes B') + \Upsilon. \tag{3.3.5c}$$

Denoting four assets by i, j, k, and l, it can be shown that since the unexplained return variation $\boldsymbol{e}_t$ is assumed to be independent of the factors, Δ is a diagonal matrix with ith diagonal element equal to the variance of the ith error term. Respectively, Ω is a $N \times N^2$ matrix of zeros except for the i, j elements where $j = (i-1)N + l$. This corresponds to the expected third moment of the idiosyncratic factors. The definition of ϒ is a bit more complicated. Similar to the other residual matrices, it consists mainly of zeros, except for the cokurtosis elements corresponding to the decomposition of the kurtosis of one asset, the cokurtosis between two assets, and the cokurtosis between three assets. (see Boudt et al. 2015, 228, 231–232 for more technical details and proofs.)

A multi-factor version of the estimators for the higher-order moment matrices involves an increase in dimensionality, which may not lead to better out-of-sample results (Martellini & Ziemann 2008, 8). There is also evidence that after the first factor, the marginal explanatory power of additional factors is relatively low (see, e.g., Connor & Korajczyk 1993; Chan et al. 1999). Thus, as in Boudt et al. (2015, 230), we use a three-factor model and compare its performance to a single-factor model, as well as to the approach based on unstructured sample estimators. We fit these two models using Principal Component Analysis (see Peterson et al. 2018, 110–111 for further technical details). Note that, alternatively, one could follow Martellini and Ziemann (2009), who extend the concept of optimal shrinkage intensities to the presence of higher-order moments. One also needs to keep in mind that there may be autocorrelation in the variance, skewness, and kurtosis. Consistent with Boudt et al. (2015, 228), we avoid making assumptions on the functional form of the dynamic linkage between future and past returns and take the rolling estimation sample approach by setting the training period and rolling window to 252 days. However, the time variation could also be taken into account parametrically, e.g., by utilizing GARCH-type of modeling of the dynamic comoments (see, e.g., Boudt et al. 2013; Ghalanos, Rossi & Urga 2015) or accounting for regime switches in the return distribution (see, e.g., Otranto 2009; Bae, Kim & Mulvey 2014). We leave these considerations for future research.

Furthermore, since portfolio moments are highly sensitive to data spikes (or outliers), we utilize the *clean boudt* argument in the PerformanceAnalytics package (see Peterson et al. 2015, 94–95) to provide more robust estimates. This ‘multivariate Winsorization’ procedure cleans returns in a time series to reduce the magnitude, but not the number or direction, of extreme observations exceeding the $1 - \alpha\%$ threshold. By default, the alpha (i.e., the probability of filtering at $1 - \text{alpha}$) is set to 0.01 (99 percent). As a measure of the extremeness of the return observation $\boldsymbol{r}_t$, its squared Mahalanobis distance (d_t^2) is used, and the 'extremeness' of the Mahalanobis distance is set to 0.001. In other words, the value of d_t^2 is limited to the 99.9 percent quantile of the χ_n^2 distribution function with n degrees of freedom.

With respect to the objective functions, we first minimize the 95 percent portfolio Expected Shortfall (ES) with a risk budget limit. In the previous subchapter, we imposed the Equal Risk Contribution (ERC) diversification constraint as a penalty. Here, we follow Bennett (2015a, 12) and minimize component contribution by setting the upper bound for the maximum allowable risk to 25 percent (i.e., no asset can contribute more than 25 percent to total portfolio risk). The estimator for the modified ES is given by:

$$ES_\alpha(\boldsymbol{\omega}) = -\boldsymbol{\omega}'\boldsymbol{\mu} + \sqrt{m_{(2)}(\boldsymbol{\omega})} \times \frac{1}{\alpha}[a_\alpha + b_\alpha k(\boldsymbol{\omega}) + c_\alpha s(\boldsymbol{\omega}) + d_\alpha s^2(\boldsymbol{\omega})], \quad (3.3.6)$$

with $a_\alpha, b_\alpha, c_\alpha$, and d_α numbers that depend on the choice of the loss level α (five percent in our case). Moreover, $s(\boldsymbol{\omega}) \equiv m_{(3)}(\boldsymbol{\omega})/m_{(2)}^{3/2}(\boldsymbol{\omega})$ and $k(\boldsymbol{\omega}) \equiv m_{(4)}(\boldsymbol{\omega})/m_{(2)}^2(\boldsymbol{\omega}) - 3$ are the portfolio skewness and excess kurtosis, respectively. (Boudt et al. 2015, 229–230.)

In the second case, we assume that the investor maximizes the expected value of the fourth-order Taylor expansion of the Constant Relative Risk Aversion (CRRA) utility function with risk aversion parameter γ. As in Boudt et al. (2015, 229), we neutralize the effect of the mean by assuming it to be zero. This leads to the following expected utility function:

$$EU_\gamma(\boldsymbol{\omega}) = -\frac{\gamma}{2} m_{(2)}(\boldsymbol{\omega}) + \frac{\gamma(\gamma+1)}{6} m_{(3)}(\boldsymbol{\omega}) - \frac{\gamma(\gamma+1)(\gamma+2)}{24} m_{(4)}(\boldsymbol{\omega}), \quad (3.3.7)$$

where $m_{(2)}(\boldsymbol{\omega})$, $m_{(3)}(\boldsymbol{\omega})$, and $m_{(4)}(\boldsymbol{\omega})$ are the second, third, and fourth portfolio moments defined in Equation (3.3.2). We test the common choice of $\gamma = 5$, but we will also consider $\gamma = 10$ as another option. To prevent situations where the portfolio's risk exposure is concentrated in the least risky assets, we impose a simple box constraint as a way of diversification. More precisely, as in Bennett (2015e, 6–8), we require that the minimum weight of any asset must be greater than or equal to zero and that the maximum weight of any asset must be less than or equal to 0.6.

Finally, in these two asset allocation problems, the optimizer searches for the portfolio weight vector $\boldsymbol{\omega}$ minimizing or maximizing the objective function, which, besides $\boldsymbol{\omega}$, depends non-linearly on the higher-order comoment matrices Σ, Φ, and Ψ under the risk diversification constraints. We utilize the *DEoptim* solver (see Ardia et al. 2010) and add the full investment and long-only constraints. The portfolios are rebalanced monthly, and the 40 constituent stocks of the MICEX index are used as the dataset.

3.4 Portfolio optimization based on ordering information

As we have discussed, modern portfolio theory produces an optimal portfolio from estimates of (uncertain) expected returns and a covariance matrix. Here, we focus on a method proposed by Almgren and Chriss (2004; 2005), where portfolio optimization is based on replacing the expected returns with sorting criteria, i.e., with information about the *order* of the expected returns but not their *values*. Since the authors' construction also allows full use of covariance information, this framework includes the Markowitzian portfolio selection as a special case and provides a generalization to a broad class of ordering information (Almgren & Chriss 2004, 1, 57). We will test how well such an approach performs if *momentum* is assumed to exist among the Russian sectoral indices.

The underlying idea behind the Almgren–Chriss (AC) framework is that for a given level of risk, an investor should prefer to hold a portfolio with a higher expected return in every scenario consistent with her beliefs. This generates a preference relation on investment portfolios, allowing one to choose between portfolios of equal levels of risk. Mathematically, a single expected return vector generates a half-space of possible expected return vectors, which have a nonnegative inner product with the given vector. For any constraint set, the construction of the efficient set and the optimal portfolio gives identical results to the classic Markowitz theory. However, unlike the mean-variance approach, this method involves, among other things, non-smooth convex optimization. (see Almgren & Chriss 2004, 3–4, 59; 2005, 2–3.)

Let us provide a simple illustration. If there are three securities S_1, S_2, S_3 and the information consists of the belief that $\rho_1 \geq \rho_2 \geq \rho_3$ where ρ_i is the expected return for S_i, then all triples of the form (r_1, r_2, r_3) where $r_1 \geq r_2 \geq r_3$ are consistent with the ordering. Therefore, it is assumed that the investor does not know the expected return vector but has distinct beliefs about the relationship between the components of the expected return vector, expressed by different sets of inequalities. Next, let us write Q for the set of all expected returns that are consistent with the given ordering and refer to them as 'consistent returns.' All available information about the expected return vector $\boldsymbol{r}$ is now listed in the form of homogenous inequality relationships, which defines the consistent cone Q, the 'coarse' preference relation, and the set of efficient portfolios. (Almgren & Chriss 2004, 4–5, 11; 2005, 21; see also Boyd & Vandenberghe 2004 for more details on the mathematics of cones.)

Indeed, if $\boldsymbol{w}$ and $\boldsymbol{v}$ are two portfolios, then, leaving aside risk limits or other constraints, $\boldsymbol{w}$ is preferable to $\boldsymbol{v}$ if the expected return of $\boldsymbol{w}$ is greater than or equal to that of $\boldsymbol{v}$ for every (concrete) expected return vector $\boldsymbol{r} = (r_1, \dots, r_n)$ consistent with the ordering (i.e., for every $\boldsymbol{r} \in Q$). Correspondingly, if $\boldsymbol{w} \succeq \boldsymbol{v}$, it is clear that $\boldsymbol{w} \cdot \boldsymbol{\rho} \geq \boldsymbol{v} \cdot \boldsymbol{\rho}$ since $\boldsymbol{\rho} \in Q$. Furthermore, if the covariance matrix of the securities in the portfolio is given by V, then for a given level of variance $\boldsymbol{\sigma}^2$ a portfolio $\boldsymbol{w}$ is efficient if $\boldsymbol{w}^{\mathrm{T}} V \boldsymbol{w} \leq \boldsymbol{\sigma}^2$ and there is no portfolio $\boldsymbol{v}$ such that $\boldsymbol{v}$ is preferable to $\boldsymbol{w}$ and $\boldsymbol{v}^{\mathrm{T}}V\boldsymbol{v} \leq \boldsymbol{\sigma}^2$. Next, let $\mathcal{M} \subset \mathbb{R}^n$ denote the set of allowable portfolio weight vectors satisfying the investor's constraints. Secondly, let $R^{\perp}$ and R define an orthogonal decomposition of the return space and of the portfolio space into two linear subspaces such that $R^{\perp} = \{\boldsymbol{r} \in \mathbb{R}^n | D\boldsymbol{r} = 0\} = Q \cap (-Q)$ and $R = (R^{\perp})^{\perp} = \{\boldsymbol{w} \in \mathbb{R}^n | \boldsymbol{w}^{\mathrm{T}}\boldsymbol{r} = 0 \text{ for all } \boldsymbol{r} \in R^{\perp}\}$. Here, R is the 'relevant' subspace defined by the investor's beliefs, and D is the $m \times n$ belief matrix containing the coefficients of the linear inequalities. Without going into further details, it can now be shown that portfolio $\boldsymbol{w}$ is efficient in the (convex) budget constraint set $\mathcal{M}$ if and only if $\mathcal{M}$ has a supporting hyperplane at $\boldsymbol{w}$ whose normal lies in both the cone Q and the hyperplane R. (see Almgren & Chriss 2004, 10–15; 2005, 7.)

Even though different sorts could be employed within this framework (e.g., higher-order sorts, partial and multiple sorts), we settle for the straightforward case of a *single complete sort*. The set of efficient portfolios for a single complete sort is precisely those that are mean-variance optimal for expected returns that are both consistent with the sort and sum to zero (i.e., there is a one-to-one correspondence between efficient portfolios and expected return vectors $\boldsymbol{r}$ such that $r_1 \geq r_2 \geq \cdots \geq r_n$ and $r_1 + \cdots + r_n = 0$). Furthermore, there are now $m = n - 1$ beliefs of the form $r_j - r_{j+1} \geq 0$ for $j = 1, \ldots, n - 1$. The belief vectors are of the form $\boldsymbol{D}_j = (0, \ldots, 0, 1, -1, 0, \ldots, 0)^{\mathrm{T}}$, and the matrix D is

$$D = \begin{pmatrix} \boldsymbol{D}_1^{\mathrm{T}} \\ \vdots \\ \boldsymbol{D}_m^{\mathrm{T}} \end{pmatrix} = \begin{pmatrix} 1 & -1 & & & \\ & 1 & -1 & & \\ & & \ddots & \ddots & \\ & & & 1 & -1 \end{pmatrix}, \tag{3.4.1}$$

where empty spaces are zeros. Hence, the consistent cone is a wedge shape in $\mathbb{R}^n$, with a 'spine' along the diagonal $(1, \ldots, 1)$ direction. (see Almgren & Chriss 2004, 12, 32, 38–39; 2005, 3–6, 11–13.)

In order to define the 'fine' version of the portfolio preference relation, it is necessary to introduce a probability measure on the set of expected returns, assessing the relative likelihood of different consistent returns being the true expected returns. Here, $\boldsymbol{\mu}$ denotes a radially symmetric probability measure on Q, and it assigns equal probability to every direction in the space of expected returns. It can be shown that when the probability distribution obeys symmetry properties, the related preference relation is entirely characterized by a linear function termed the *centroid*. This function is the center of mass of the set Q under the measure $\boldsymbol{\mu}$ as defined by the integral $\boldsymbol{c} = \int_{\boldsymbol{r} \in Q} \boldsymbol{r} \, \mathrm{d}\boldsymbol{\mu}$. Therefore, the whole problem of calculating efficient or optimal portfolios can be reduced to manipulations involving the centroid vector $\boldsymbol{c}$. (see Almgren & Chriss 2004, 5–7, 28–32; 2005, 13–14.)

Indeed, finding the preferable portfolio relative to a maximum risk constraint is equivalent to finding the maximum of $\boldsymbol{c}$ on the set $\mathcal{M}$ of portfolios respecting this constraint. Hence, we must solve the following linear program with quadratic constraints:

$$\max_{\boldsymbol{w}} \ \boldsymbol{w} \cdot \boldsymbol{c} \quad \text{subject to} \quad \boldsymbol{w} \cdot V \cdot \boldsymbol{w} \leq \boldsymbol{\sigma}^2, \tag{3.4.2}$$

where the vector of portfolio weights $\boldsymbol{w}$ can be considered as a sum $\boldsymbol{w} = \sum_{i=1}^{n-1} x_i D_i + x_n(1, \dots, 1)^{\mathrm{T}}$ with $x_1, \dots, x_n$ being real numbers, and $\boldsymbol{c} \cdot \boldsymbol{w}$ is the scalar quantity over $\boldsymbol{w} \in \mathcal{M}$ to be maximized. Furthermore, we use the direct calculation method to estimate the centroid vector. More precisely, for a single complete sort of n assets, the jth component of $\boldsymbol{c}$ can be analytically approximated as follows:

$$c_{j,n} = N^{-1}\left(\frac{n+1-j-\alpha}{n-2\alpha+1}\right), \qquad \alpha = A - Bn^{-\beta}, \tag{3.4.3}$$

where $N^{-1}(\cdot)$ is the inverse cumulative normal distribution and $A = 0.4424$, $B = 0.1185$ and $\beta = 0.21$. This gives centroid components with maximum fractional error less than one-half of percent when n is very small, decreasing rapidly as n increases. We then scale the estimated centroid vector according to the median of the asset mean returns and reorder it such that the highest (smallest) centroid value is assigned to the asset index with the highest (lowest) expected return. (Almgren & Chriss 2004, 14, 21, 59–61; 2005, 8–9, 15–16; see also Peterson et al. 2018, 10 for further technical details.)

It is worth emphasizing that the centroid vector for a complete sort of assets overweights very high and very low ranked assets while underweighting the middle. In other words, it can be shown that the centroid portfolio curves up at the ends, placing a larger weight on the differences at the ends than on the differences in the middle of the portfolio. This is because distributions tend to have long tails, so two neighboring samples near the endpoints are likely to be more different from each other than two samples near the middle. On the other hand, Almgren and Chriss (2004, 33–34, 50–57) also demonstrate that their portfolio construction methodology is robust to fairly large levels of information degradation. Naturally, this is important since ranking errors are inevitable in real situations due to imperfect knowledge of the exact order of future expected returns.

As mentioned earlier, our modified AC trading strategy assumes that the assets with the highest return will continue to outperform. Hence, we follow Bennett (2015c) and define a factor-function where the ‘momentum period’ K is set to 21, and the sort parameter on date $t-1$ is the positive of the cumulative return from date $t-K$ to $t-1$. As daily data is used, this means that the algorithm ranks the sectoral indices based on past returns over monthly periods. Such an approach is in line with Moskowitz and Grinblatt (1999, 1249–1251, 1286–1287), who find that industry momentum is strongest at the one-month horizon. They also argue that industry-based strategies are more profitable than individual stock momentum strategies.

To conduct an out-of-sample backtest, we impose both the long-only and full investment constraints, set the training period to 126 days (half a trading year), and rebalance the portfolios monthly. As before, the *DEoptim* (see Ardia et al. 2010; Mullen et al. 2011; Ardia et al. 2020) is used for optimization. Furthermore, we compare the performance of the AC procedure to three competing variants. The first is the equal-weight strategy, and the second is the classical quadratic utility maximization strategy, where sample estimates are used for the first and second moments, and the risk aversion parameter, λ, is set to

0.5. The third is a ranking strategy based on Meucci's (2008b) framework of fully flexible views.

Let us briefly outline the implementation of the last-mentioned approach. As discussed in Subchapter 2.3, the posterior distribution can be derived by assuming that the market for N assets is modeled by a sample of the securities' (random) returns $\mathbf{X}$ following an arbitrary a priori joint distribution, $f_{\mathbf{M}}$. Here, we first retrieve the empirical distribution from historical data and then use Monte Carlo to obtain simulated values for this reference model. Setting $J = 1000$, a matrix $\mathcal{M}$ of dimension $(J \times N)$ is thus created where the columns are the marginal prior distributions, and the rows are the simulated outcomes for the market factors. Next, a probability p_j is associated with each of these outcomes $\mathcal{M}_{j,\cdot}$. We follow the most straightforward way and set the J probabilities equal to $1/J$ (i.e., we consider each Monte Carlo draw to be equally likely). Hence, the $(J \times 1)$ probability vector $\boldsymbol{p}$ has as elements the reciprocal of the size of the Monte Carlo simulation. (Pfaff 2016, 294; see also Meucci 2008b, 8.)

To express views on the relative order of asset returns, we set $\widetilde{m}\{V_1\} \geq \widetilde{m}\{V_2\} \geq \cdots \geq \widetilde{m}\{V_K\}$. When the location measure $\widetilde{m}\{V_k\}$ is the expectation, this translates into the following set of linear constraints:

$$\begin{gathered} \sum_{j=1}^{J} \tilde{p}_j \left(\mathcal{V}_{j,1} - \mathcal{V}_{j,2}\right) \geq 0 \\ \vdots \\ \sum_{j=1}^{J} \tilde{p}_j \left(\mathcal{V}_{j,K-1} - \mathcal{V}_{j,K}\right) \geq 0\,. \end{gathered} \tag{3.4.4}$$

Denoting the function values by $\boldsymbol{g}$, the implied distribution of these views is then approximated by the market simulations $\mathcal{M}$ according to $\mathcal{V}_{j,k} = g_k\left(\mathcal{M}_{j,1}, \dots, \mathcal{M}_{j,N}\right)$ for $k = 1, \dots, K$ and $j = i, \dots, J$. (Meucci 2008b, 4, 19.)

Once the two panels, $\mathcal{M}$ and $\mathcal{V}$, have been created, they are combined in three steps to retrieve the posterior market distribution obeying the views. First, the views are expressed in terms of a set of linear inequality constraints, $\boldsymbol{a}_{lower} \leq A\overline{\boldsymbol{p}} \leq \boldsymbol{a}_{upper}$, where the lower and upper bounds $(\boldsymbol{a}_{lower}, \boldsymbol{a}_{upper})$ and the matrix A are derived from $\mathcal{V}$, and the probability vector $\overline{\boldsymbol{p}}$ is now treated as the objective variable in the ensuing optimization. In the second step, the relative entropy (or Kullback–Leibler divergence) is minimized. We define the discrete objective function as

$$\mathrm{RE}(\overline{\boldsymbol{p}}, \boldsymbol{p}) = \sum_{j=1}^{J} \bar{p}_j \left[\log(\bar{p}_j) - \log(p_j)\right]. \tag{3.4.5}$$

Hence, the probabilities of the posterior distribution under the assumption of perfect foresight can be obtained by evaluating

$$\overline{\boldsymbol{p}} = \underset{\boldsymbol{a}_{lower} \leq A\overline{x} \leq \boldsymbol{a}_{upper}}{\arg\min} \mathrm{RE}(\boldsymbol{x}, \boldsymbol{p}). \tag{3.4.6}$$

As shown in Meucci (2008b, 22), the dual form of the above optimization can be derived from the Lagrangian function. This results in a linearly constrained convex program with the size of the objective variable equal to the number of views K. Thirdly, we determine the empirical confidence-weighted posterior distribution $(\mathcal{M}, \boldsymbol{p}_c)$ according to $\boldsymbol{p}_c = (1-c)\boldsymbol{p} + c\overline{\boldsymbol{p}}$, where the confidence level in the views, c, is set to 0.5. The constructed posterior distribution is then be utilized in a portfolio optimization problem, where the objective is again to maximize return while penalizing variance. As such, this approach satisfies the views and changes with the empirical data $(\widehat{\boldsymbol{\mu}}, \widehat{\boldsymbol{\sigma}}^2)$. (Pfaff 2016, 294–295; Bennett 2015abc.) To execute the entropy pooling program, we utilize the algorithms contained in the *PortfolioAnalytics* package (see Peterson et al. 2018, 48–50, 74–75).

3.5 Tactical asset allocation within the Black–Litterman framework

This subchapter employs the Copula Opinion Pooling (COP) and Entropy Pooling (EP) concepts and outlines two ways to obtain portfolio allocations from a set of quantitatively derived signals. The theoretical background behind these strategies draws on the works of Meucci (2006; 2008ab), whereas our R implementation follows the approach proposed by Pfaff (2016). In the first part, we apply the Black–Litterman (BL) framework to the one-step-ahead forecasts for the constituents of the Transport sector derived from a Vector Error Correction Model (VECM). The estimated parameters of the posterior distributions are then used to obtain a weight vector that maximizes the portfolio's Sharpe ratio. The application involves a backtest, where in addition to the BL model, portfolio weights based upon the prior distribution and a naïve allocation are computed. The resulting portfolios' equities are then contrasted with each other. Next, we continue by employing the latest return forecasts in the COP model and compare the obtained portfolio allocations with the BL model. This is because, in the original BL model, the views are modeled using the normality assumption, implying independence between the expressed views. Arguably, it is not sensible to assume that return forecasts for various assets are formed independently of each other, especially when these are deduced from a multivariate statistical model. The BL model also assumes that the returns follow a multivariate normal distribution, although the stylized facts of financial returns tend to conflict with this assumption. (see Pfaff 2016, 274, 289, 292, 307, 313.)

Unfortunately, views on the dispersion or volatility of assets are not feasible in the COP model. Therefore, in the second part, we combine the EP model with short-term views derived from GARCH(1,1) models and use the constituents of the MICEX index as our dataset. Such an approach is in line with Hayo and Kutan (2005, 378), who note that a GARCH(1,1) model is a specification that works well for both bond and stock markets in Russia with regard to capturing ARCH effects. Similarly, Abrosimova et al.

(2005, 12–13) find that GARCH(1,1) fits the Russian stock market data best and provides the most accurate one-day-ahead forecasts in comparison to the other volatility models, such as GARCH-M, AGARCH, and EGARCH.

Let us begin by explaining the basic principles of the COP approach. The starting point is the arbitrary prior market distribution, expressed in terms of its probability distribution function (pdf), i.e., $\mathbf{M} \sim f_{\mathbf{M}}$. It represents the returns on a set of N assets, and investors twist this 'prior benchmark' according to their subjective views. These $K \leq N$ views on linear combinations of the market $\mathbf{M}$ are represented by a $K \times N$ dimensional 'pick' matrix $\mathbf{P}$, and thus the generic kth row of the pick matrix determines the weights of the kth views. One may then choose $N - K$ additional complementary linear combinations on which no view is expressed. These complementary directions are compacted as rows of a $(N - K) \times N$ matrix $\mathbf{P}^{\perp}$, and the resulting invertible matrix $\overline{\mathbf{P}} \equiv \binom{\mathbf{P}}{\mathbf{P}^{\perp}}$ defines the view-adjusted market coordinates. In other words, the N-dimensional random vector $\mathbf{V} \equiv \overline{\mathbf{P}}\mathbf{M}$ is fully equivalent to the market $\mathbf{M}$. (Meucci 2006, 3.)

Furthermore, as the views correspond to statements on the first K entries of $\mathbf{V}$, each of these statements can be expressed in terms of a Cumulative Distribution Function (CDF):

$$\hat{F}_k(v) \equiv \mathbb{P}_{subj}\{V_k \leq v\}, \qquad k = 1, \dots, K. \tag{3.5.1}$$

To perform marginalizations, we represent a distribution for each view by its CDF:

$$F_k(v) \equiv \mathbb{P}_{prior}\{V_k \leq v\}, \qquad k = 1, \dots, K. \tag{3.5.2}$$

Since the views (3.5.1) are typically different from the respective market-implied distributions (3.5.2), opinion pooling techniques are needed to resolve such a dichotomy. The posterior CDF can thus be defined as a weighted average:

$$\tilde{F}_k \equiv c_k \hat{F}_k + (1 - c_k) F_k, \qquad k = 1, \dots, K, \tag{3.5.3}$$

where the weight c_k, $k = 1, \dots, K$ represents the confidence in the respective view. (Meucci 2006, 3–4.) We set $c_k = 0.5$, meaning that we assume equal confidence in all the views. Note that this parameter is similar to the diagonal elements of the uncertainty matrix in the BL model.

Formula (3.5.3) defines the posterior marginal distribution of each view. To perform the copula factorization, we inherit the copula of the posterior distribution of the views from the (market-implied) prior copula as follows:

$$\begin{pmatrix} C_1 \\ \vdots \\ C_K \end{pmatrix} \stackrel{d}{=} \begin{pmatrix} F_1(V_1) \\ \vdots \\ F_K(V_K) \end{pmatrix}. \tag{3.5.4}$$

Hence, the posterior joint distribution of the views is defined by the quantile inversions:

$$\begin{pmatrix} V_1 \\ \vdots \\ V_K \end{pmatrix} \stackrel{d}{=} \begin{pmatrix} \tilde{F}_1^{-1}(C_1) \\ \vdots \\ \tilde{F}_K^{-1}(C_K) \end{pmatrix}. \tag{3.5.5}$$

Finally, in order to determine the market's posterior distribution, $\tilde{f}_{\mathbf{M}}$, $\mathbf{V} \equiv \overline{\mathbf{P}}\mathbf{M}$ is applied backward, and the views are rotated back into the market coordinates. A simple matrix multiplication leads to $\mathbf{M} \stackrel{d}{=} \overline{\mathbf{P}}^{-1}\mathbf{V}$, where the first K entries of $\mathbf{V}$ follow the posterior distribution (3.5.5), and the remaining entries are left unaltered. (see Meucci 2006, 4, 7–9 for more technical details.) As in BL, the COP posterior distribution represents a subtle, consistent twist of the original prior distribution. Hence, it should give rise to sensible allocations once fed into an optimization algorithm.

In R, the above steps can be executed by utilizing the *BLCOP* package (see Gochez et al. 2020). It relies on the fact that the multivariate distribution $\mathbf{M}$ can be represented by means of a large number of Monte Carlo simulations (i.e., by a $J \times N$ panel $\mathcal{M} \Leftrightarrow \mathbf{M} \sim f_{\mathbf{M}}$, where each N-dimensional row represents one of the J simulated joint scenarios for the market variable $\mathbf{M}$). Thus, the posterior marginal CDFs in (3.5.3) can be approximated with high accuracy by their empirical counterparts, (3.5.5) can be retrieved by means of interpolation between given grid points, and the joint posterior simulated realizations of the market distribution can be swiftly recovered. In line with Pfaff (2016, 313), we set the simulation size to 10,000 random draws and assume a *multivariate skewed Student's* t distribution (see Azzalini & Capitanio 2014, 176) for the asset returns. The parameters of this distribution are estimated by the maximum log-likelihood estimation method contained in the *fMultivar* package (see Würtz, Setz & Chalabi 2020, 4).

We can now proceed to the generation of views. As is well known, Vector Error Correction Models (VECMs) unify two branches of research, namely Vector Autoregressive (VAR) models and the concept of co-integration introduced by Granger (1981) and Engle and Granger (1987). In its basic form, a VAR consists of a set of K endogenous variables $\boldsymbol{y}_t = (y_{1t}, \dots, y_{kt}, \dots, y_{Kt})$ for $k = 1, \dots, K$. By omitting deterministic regressors (e.g., constants and dummies), the vector autoregression model of order p can be defined as

$$\boldsymbol{y}_t = A_1\boldsymbol{y}_{t-1} + \cdots + A_p\boldsymbol{y}_{t-p} + \boldsymbol{u}_t, \tag{3.5.6}$$

where the A_i are $(K \times K)$ coefficient matrices for $i = 1, \dots, p$ and $\boldsymbol{u}_i$ is a K-dimensional white noise process with $\mathrm{E}(\boldsymbol{u}_t) = \mathbf{0}$ and with time-invariant positive definitive covariance matrix $\mathrm{E}(\boldsymbol{u}_t\boldsymbol{u}_t') = \Sigma_{\boldsymbol{u}}$. Once the coefficients of a VAR(p) process have been estimated by least squares applied separately to each of the equations, forecasts for horizons $h \geq 1$ can be generated recursively from

$$\boldsymbol{y}_{T+h|T} = A_1\boldsymbol{y}_{T+h-1|T} + \cdots + A_p\boldsymbol{y}_{T+h-p|T}, \tag{3.5.7}$$

where $\boldsymbol{y}_{T+j|T} = \boldsymbol{y}_{T+j}$ for $j \leq 0$. The variance-covariance matrix of the forecast errors is a function of $\Sigma_{\boldsymbol{u}}$ and Φ_s, where the Φ_s are the coefficient matrices of the Wold moving average representation of a stable VAR(p) process. (see Pfaff 2016, 284–286.)

From the above VAR(p) process, we specify the vector error correction by using the following *transitory* form:

$$\Delta\boldsymbol{y}_t = \boldsymbol{\alpha}\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{y}_{t-1} + \Gamma_1\Delta\boldsymbol{y}_{t-1} + \cdots + \Gamma_{p-1}\boldsymbol{y}_{t-p+1} + \boldsymbol{u}_t, \tag{3.5.8a}$$

$$\Gamma_i = -\left(A_{i+1} + \cdots + A_p\right), \quad \text{for } i = 1, \dots, p-1, \tag{3.5.8b}$$

$$\Pi = \boldsymbol{\alpha}\boldsymbol{\beta}^{\mathrm{T}} = -\left(I - A_1 - \cdots - A_p\right). \tag{3.5.8c}$$

Here, I is the $(K \times K)$ identity matrix, and the levels of the components in $\boldsymbol{y}_t$ enter lagged by one period. The Γ_i matrices measure transitory effects, and in the case of co-integration, the matrix $\Pi = \boldsymbol{\alpha}\boldsymbol{\beta}^{\mathrm{T}}$ is of reduced rank. The dimensions of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are $K \times r$, and r is the co-integration rank, denoting how many long-run relationships exist between the elements of $\boldsymbol{y}_t$. The matrix $\boldsymbol{\alpha}$ is the loading matrix, and the coefficients of the dynamic interdependencies are contained in $\boldsymbol{\beta}$. (Pfaff 2008, 78–80; 2016, 287–288; see also the monographs of Banerjee et al. 1993; Hamilton 1994; Hendry 1995; Johansen 1995; Lütkepohl 2004; 2005 for in-depth expositions of these models.)

In order to assess the degree of integration of a time series, we employ the (augmented) Dickey-Fuller (ADF) and Elliot-Rothenberg-Stock (ERS) unit root/stationarity tests included in the *urca* package (see Pfaff, Zivot & Stigler 2016, 10, 44, 46). We then carry out the estimation and testing of the cointegration rank in a VECM by following the procedure proposed, among others, by Johansen and Juselius (1990; 1992) and Johansen (1995) and report the maximal eigenvalue test statistics. In brief, this method uses canonical correlation analysis as a tool to reduce the information content of T observations in the K-dimensional space to a lower-dimensional one of r cointegrating vectors (see Pfaff 2008, 81–82). Finally, to compute the point forecasts with associated confidence bands, we express the estimated VECM in its level-VAR form by using the transformation function included in the *vars* package (see Pfaff & Stigler 2018, 25–27, 48–49).

Let us next turn to the EP model. Its principles were already introduced in the previous subchapter, and here we only slightly modify that approach. More precisely, we seek the solution to the tangency portfolio, assume a skewed Student's t distribution as prior market distribution, and express a view on volatility as $\tilde{\sigma}\{V_k\} \gtreqless \sigma_k$. Since $\tilde{\sigma}\{V_k\}$ is the standard deviation, we can define the equality constraints with respect to the views as follows:

$$\sum_{j=1}^{J} \tilde{p}_j V_{j,k}^2 = \widehat{m}_k^2 + \sigma_k^2, \tag{3.5.9}$$

where $\widehat{m}_k$ denotes the first (sample) moment of the kth asset. We utilize the *fGarch* package (see Würtz et al. 2020) to estimate the GARCH models and compute the one-step-ahead predictions for the conditional volatilities for each asset. Furthermore, we split our full data sample into two subperiods and use the data points in the latter subperiod in a portfolio backtest, where optimal asset allocations are determined based on the EP posterior, the market distribution, as well as according to a multivariate normal distribution assumption. The outcome of these three portfolio strategies is then contrasted. Finally, the *fPortfolio* (see Würtz et al. 2015b) and *PerformanceAnalytics* (see Peterson et al. 2015) packages are employed for conducting the portfolio optimization and evaluation of the backtest results. (see Meucci 2008b, 4, 19; Pfaff 2016, 318–323 for more details.)

3.6 Wealth protection with return forecasting

The main purpose of different protection strategies is to stop the deterioration of financial wealth positions, and therefore it is crucially important for any equity investor to be aware of them, especially during financial crises. A typical approach has been to utilize an option-based insurance strategy, which protects a portfolio from market declines without losing the opportunity to participate in market rallies, similar to a protective put (see, e.g., Kritzman 1986 & 1990; Mason et al. 1995; Lyuu 2004, 468–469 for more discussions). In contrast, Pfaff (2007; 2016, 326–334) proposes a novel way to construct a TAA-related wealth protection strategy where one does not need to resort to derivatives. Analogously to the previous approaches, the idea here is to specify a portfolio floor value that may not be violated. The difference between the current wealth and this floor is then used as a risk buffer in which the risk budget is allocated between the assets in proportion to their downside risk. Such a constraint can be expressed as a linear program with return maximization as its target. Hence, we now need to combine three pieces, namely a forecasting model, a risk model, and a linear program.

Let us start with the first piece. In line with the ‘principle of parsimony,’ we wish to utilize a model with as few parameters as possible but which gives an adequate representation of the data at hand. Secondly, we recall from Fildes and Makridakis (1995) that although complex models may give a better fit than simpler models, the resulting (short-term) forecasts need not be more accurate. Thirdly, as discussed in Chatfield and Xing (2019, 60, 141), the differences in accuracy for different methods when averaged over many series tend to be relatively small compared with the large differences in accuracy that may arise when different methods are applied to individual series. Based on this reasoning, we decide to take the return expectations as the one-step-ahead point forecasts from an Autoregressive Integrated Moving Average (ARIMA) model.

Following the same notation as in Hyndman and Athanasopoulos (2014, 225–240), a non-seasonal ARIMA(p, d, q) model can be written as

$$y'_t = c + \phi_1 y'_{t-1} + \cdots + \phi_p y'_{t-p} + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t, \tag{3.6.1}$$

where y'_t is the differenced series (i.e., $y'_t = y_t - y_{t-1}$). If the differenced data is not stationary, one may need to difference it a second time to obtain a stationary series (i.e., $y''_t = y'_t - y'_{t-1}$). In (3.6.1), the ‘predictors’ on the right-hand side include both lagged values of y_t and lagged errors. In addition, c denotes the constant and ε_t is white noise. This equation can be written in backshift notation as

$$(1 - \phi_1 B - \cdots - \phi_p B^p)(1 - B)^d y_t = c + (1 + \theta_1 B + \cdots + \theta_q B^q)\varepsilon_t, \tag{3.6.2}$$

or equivalently as

$$\begin{aligned}(1 - \phi_1 B - \cdots - \phi_p B^p)(1 - B)^d (y_t - \mu t^d / d!)\\ = (1 + \theta_1 B + \cdots + \theta_q B^q)\varepsilon_t,\end{aligned} \tag{3.6.3}$$

where $c = \mu(1 - \phi_1 - \cdots - \phi_p)$ and μ is the mean of $(1 - B)^d y_t$. Thus, the inclusion of a constant in a non-stationary ARIMA model is equivalent to inducing a polynomial trend of order d in the forecast function. Note that when utilizing the *forecast* package (see Hyndman et al. 2018, 10–13), the parameterization of (3.6.3) is used, and $c = \mu = 0$ by default when $d > 0$. Moreover, as the fitting method, this package utilizes conditional-sum-of-squares to find starting values and then maximum likelihood estimation. The AR and MA order and the order of integration can also be determined automatically using the Akaike information criterion (see Hyndman & Khandakar 2008, 15–18).

The recursive forecasting procedure can be outlined as follows: In the first step, the algorithm expands the ARIMA equation so that y_t is on the left-hand side, and all other terms are on the right. Then, it rewrites the equation by replacing t with $T + h$. In the third step, it replaces (on the right-hand side of the equation) future observations with their forecasts, future errors with zero, and past errors with the corresponding residuals. Beginning with $h = 1$, these steps can be further repeated for $h = 2,3, \ldots$ until all forecasts $\hat{y}_{T+h|T}$ have been calculated. (see Hyndman and Athanasopoulos 2014, 239–240 for an illustration.) Eventually, the return expectation can be formed as the difference between the latest log value and the generated point forecast.

The second piece consists of a market price risk model. We use the Expected Shortfall (ES) for a confidence level of 95 percent and derive this downside risk measure from the Generalized Pareto Distribution (GPD). More precisely, the GPD is defined as

$$H(y) = 1 - \left(1 + \frac{\xi y}{\tilde{\sigma}}\right)^{-1/\xi}, \tag{3.6.4}$$

for $y : y > 0$ and $\tilde{\sigma} = \sigma + \xi(u - \mu)$. Here, μ is the location, σ the scale, ξ the shape parameter, and u is the threshold value. We estimate these parameters by considering the 30 largest losses and applying the Probability-Weighted Moments (PWMs) method contained in the *evir* package (see Pfaff et al. 2018, 9–10, 26). The Extreme Value Theory-based risk measures VaR and ES can then be inferred directly from the GPD as follows:

$$\mathrm{VaR}_\alpha = q_\alpha(F) = u + \frac{\tilde{\sigma}}{\xi}\left(\left(\frac{1-\alpha}{\bar{F}(u)}\right)^{-\xi} - 1\right), \tag{3.6.5a}$$

$$\mathrm{ES}_\alpha = \frac{1}{1-\alpha}\int_\alpha^1 q_x(F)\,\mathrm{d}x = \frac{\mathrm{VaR}_\alpha}{1-\xi} + \frac{\tilde{\sigma} - \xi u}{1-\xi}, \tag{3.6.5b}$$

where $\bar{F}(u)$ denotes the number of non-exceedances relative to the sample size. Since the riskiness is determined for each loss series alone, the weighted sum thereof forms an upper bound for the overall portfolio downside risk. In other words, a perfect concordance of the asset losses is assumed. Next, given that losses are expressed as positive numbers, we calculate the sign factors containing the correct trade/loss direction from the return forecasts. The factors are then used to alter the signs of the actual returns such that these comply with the relevant losses from long positions. (cf. Pfaff 2016, 92–94, 329–330.)

The third part is the formulation of the linear program, which is specified with four arguments. In the first argument, the return expectations are taken as the parameters of the objective function to be maximized. The second argument holds the downside risk ES estimates derived from the GPD. The third argument contains a buffer for the risk budget such that the portfolio's downside risk as a weighted sum of its components shall not exceed this limit. The last argument is an upper bound on the portfolio weights, and we set this constraint to 40 percent in order to circumvent concentrated portfolio solutions. The weights themselves are treated as non-negative real numbers, and an additional budget constraint is defined such that the sum of weights shall not exceed unity. The resulting linear programming problem is then solved by employing the algorithm contained in the *Rglpk* package (see Theussl et al. 2019, 3–5). (see also Pfaff 2016, 330–331 for more technical details.)

Finally, the previously created objects are used to conduct a recursive backtest, where the sectoral indices are used as the dataset, and the portfolios are rebalanced weekly. As in Pfaff (2016, 332–333), we initialize the portfolio equity to 100 monetary units and assume that the funds can be invested at a money market rate of 1 percent per annum. We determine the 90 percent current wealth level within the portfolio simulation and compare it to the portfolio floor to be protected. The optimization is carried out only if the former is greater than the latter. Otherwise, the remaining funds earn interest at the money market rate until the initial floor level is exceeded again. Given that the forecasts for all assets can be negative at the same time, and hence a long-only investor would lose from exposure to any of these assets, the wealth would also earn interest only. In Subchapter 4.5, we express the proceeds from the current allocation in terms of the portfolio wealth and compare the outcome of this simulated wealth trajectory with the solution of an equal-weighted strategy.

4 EMPIRICAL FINDINGS AND CONCLUSIONS

Theoretical elegance alone does not necessarily deliver improved investment performance. Therefore, we now analyze whether our portfolio construction and investment methods would have delivered acceptable performance results in practice. Finally, having compared our findings with the previous studies, we provide suggestions for future research.

4.1 Backtest analysis of classical and robustified mean-risk portfolios

We start this subchapter by studying the out-of-sample performance of classical minimum variance and maximum return portfolios in the presence of transaction costs. As we have discussed, both of these approaches can be modified in several ways. Thus, before moving on to inspecting whether robust covariance estimators can improve the performance of minimum variance portfolios, we analyze the impact of a risk budget constraint on a portfolio maximizing the mean return per unit Expected Shortfall (i.e., the STARR ratio). Finally, we compare the portfolios produced by robust optimization techniques to classical Markowitz portfolios. Approaches to future research are proposed at the end of each application. As our dataset, we use the total return indices of the 40 constituents of the MICEX index.

Table 4.1 shows the performance and risk metrics of the variance minimization (*min var*) and return maximization (*max ret*) strategies in addition to the maximum STARR ratio (*max ret per unit ETL*) portfolio without and with the risk budget constraint. More precisely, the objective of the fourth portfolio is to maximize the STARR ratio such that each component contributes equally to portfolio downside risk as measured by the modified Expected Tail Loss (Expected Shortfall or Conditional Value-at-Risk). Thus, to construct a risk budget, we add an Equal Risk Contribution (ERC) constraint where $\%C_i\mathrm{ETL}(w) = 1/N \ (i = 1, \dots, N)$. The MICEX gross total return index serves as our benchmark, and the preferable values are highlighted in bold. The optimizations are run with monthly rebalancing, and the first rebalance date is on 3 October 2011. The width (i.e., the number of periods) of the rolling window is set to null, meaning that all available past data is given to the rebalance function. Finally, the proportional transaction cost (*TC*) value is set to 50 basis points (0.5 percent). It is worth mentioning that we also experimented with weekly rebalancing and noticed that the risk-adjusted returns (measured by the annualized Sharpe ratios) were less attractive. Hence, due to estimation errors, one should prefer less frequent rebalancing. This conclusion remained unchanged when the transaction costs were not considered.

Although most of the metrics in Table 4.1 are self-explanatory or have been introduced earlier, a few clarifications should be made. Firstly, the annualized Sharpe ratio is computed by dividing the annualized mean excess return by the annualized standard deviation of excess return, i.e., $\left(\sqrt[n]{prod(1+R_a)^{scale}}-1\right)/\left(\sqrt{scale}\cdot\sqrt{\sigma}\right)$, where the daily scale is set to 252, and n is the total number of periods for which there are observations (see Peterson et al. 2017, 175). The assumed risk-free rate is denoted by R_f. Secondly, in addition to the annualized Sharpe ratio, we report the Sortino ratio of performance over downside risk. The Sortino ratio follows the same principle as the Sharpe ratio but instead of discounting a portfolio's return against all its volatility (standard deviation of returns), only its downside semivariance is used as the measure of risk. In other words, if a portfolio is volatile on the way up, it should not be penalized. Here, the Minimum Acceptable Return (MAR) is assumed to be zero percent. Thirdly, we report the Information Ratio (IR), which is calculated as the active premium divided by the tracking error. IR relates the degree to which a portfolio has outperformed the benchmark (here, the MICEX index) to the consistency with which the portfolio has beaten the benchmark. A higher IR means that the portfolio has consistently outperformed its benchmark, making it less likely that the relative benchmark returns are by chance.

Fourthly, we perform linear regression analysis and calculate the adjusted R-squared estimate, which measures how much of a portfolio's performance is correlated with its benchmark. A low R-squared score shows that the portfolio has done well independent of the benchmark (i.e., something else can be attributed to the portfolio's returns). (see Sharpe 1994; Sortino & Price 1994; Peterson et al. 2015; Katz 2019.) The maximum drawdown, in turn, is a risk measure, which tells the worst cumulative loss ever sustained by the portfolio. In order to calculate the average monthly turnover, we follow the method proposed by Kipnis (2016). This turnover quantity can be interpreted as the average amount of wealth traded in each rebalancing period (Boudt et al. 2013, 64). Finally, as some of the modified VaR and ES risk measure calculations produced unreliable results (inverse risk or risk over 100 percent), we cleaned extreme observations in each time series using the method proposed by Boudt et al. (2008). By utilizing data cleaning, we were able to produce more robust modified risk estimates. However, the reported modified risk estimates may not be as conservative as one would get with uncleaned series.

Overall, we can see from these backtesting results that, even with the inclusion of reasonably high transaction costs, the minimum variance portfolio outperforms the MICEX index in terms of annualized Sharpe ratios (0.82 and 0.61, respectively). In contrast, the maximum return portfolio, which is at the opposite end of the efficient frontier, produces the lowest annualized return (0.09) and highest standard deviation (0.29), leading to the lowest Sharpe ratio of 0.29 (the figures have been rounded for better readability). Such a result is not surprising, given that the arithmetic mean (an estimator for the location of a population) is more prone to estimation error and more sensitive to extreme observations than the sample covariance estimator (see, e.g., Jorion 1986). Nevertheless, when the return is maximized per unit ES, the annualized return increases to 0.16, the standard deviation decreases to 0.25, leading to the Sharpe ratio of 0.65. A further significant improvement to the portfolio's performance can be achieved by imposing the risk-budget constraint. This combination produces the highest annualized return (0.19), lowest standard deviation (0.18), and thus the highest Sharpe ratio of 1.06.

Furthermore, the outperformance order based on the Sortino ratio is in line with the annualized Sharpe ratio, i.e., the maximum STARR ratio portfolio with the risk-budget constraint has the highest score of 0.10, whereas the maximum return portfolio (0.04) is the only one underperforming the benchmark index (0.06). However, although the risk-constrained portfolio also has a higher information ratio (1.04) than the minimum variance portfolio (0.37), the latter is less correlated with the MICEX index in terms of R^2 (0.87 and 0.63, respectively). Moreover, the minimum variance portfolio is the only one with positive skewness (2.17), whereas the maximum return portfolio has the most negative skewness of -1.31. Each portfolio's return distribution is leptokurtic, and the MICEX index (maximum return portfolio) has the lowest (highest) kurtosis of 5.28 (239.27).

In terms of riskiness, the maximum return portfolio has the most severe maximum drawdown of 0.61 and the largest historical ES of -0.34. The risk-constrained portfolio, in turn, has the second smallest maximum drawdown of 0.28 after the benchmark index (0.23) and the smallest historical VaR of -0.016. Based on the other risk metrics, namely historical ES (-0.026), modified VaR (-0.017), and modified ES (-0.025), both the minimum variance portfolio and the risk-constrained portfolio are equally low-risk. On the other hand, the risk-constrained portfolio has a lower average monthly turnover (0.07) than the minimum variance portfolio (0.13). Surprisingly, the maximum STARR ratio portfolio has the highest turnover of 1.26.

Figure 4.1 presents a three-panel performance summary chart covering the out-of-sample backtesting period from October 2011 to September 2017. The top panel shows the cumulative returns of the four competing strategies versus the MICEX index. Note that the value of the initial investment is set to zero. The second panel, in turn, shows the daily returns of the benchmark, and the third panel is a drawdown chart.

As we can see from the temporal evolution of accumulated returns, the maximum return portfolio performs the worst, and its cumulative returns stay negative until October 2015. The classic variance minimization approach eventually outperforms the MICEX index, although this happens only after the beginning of 2016. The cumulative returns of this portfolio and the maximum STARR ratio portfolio are rather close to each other (1.46 and 1.39, respectively), but the accumulated returns of the latter are consistently negative until the second half of 2014. In contrast, the maximum STARR ratio portfolio with the ERC constraint follows the benchmark index more closely since the beginning of the sample period, and the outperformance begins already during the spring of 2015. Of course, this was right after the highly volatile period in the Russian market caused, inter alia, by the ruble crash. Overall, the periods where one strategy is outperforming the other are relatively long and imply the possibility of applying market timing strategies on top of these allocations.

Furthermore, by examining the synchronicity of the loss periods, it can be observed that the return maximization strategy systematically produces the most severe drawdowns, followed by the maximum STARR ratio portfolio. At the beginning of the Ukrainian crisis in March-April 2014, the drawdown of the maximum return portfolio reaches the level of -0.6, whereas the drawdowns of the minimum variance and the risk-constrained maximum STARR portfolio increase sharply to roughly -0.3. However, after the sudden shock, the values decrease rather steadily and stay above the -0.2 level. The calmest period in terms of drawdown occurs during the first half of 2015 and the end of 2016, when the drawdowns remain above the -0.1 level, except in the case of the maximum return portfolio. Since October 2016, the drawdowns of each strategy have moved in line with the MICEX index.

Table 4.1 Performance metrics of the classical mean-risk portfolios with a proportional transaction cost constraint and monthly rebalancing

	MICEX index	*Min var (TC=50 bps)*	*Max ret (TC=50 bps)*	*Max ret per unit ETL (TC=50 bps)*	*Max ret per unit ETL with ETL equal contribution to risk (TC=50 bps)*
Cumulative Return	0.927	1.457	0.622	1.390	**1.799**
Annualized Return	0.118	0.165	0.086	0.159	**0.191**
Annualized Std Dev	0.195	0.201	0.292	0.247	**0.180**
Annualized Sharpe (Rf=0.00)	0.605	0.821	0.293	0.645	**1.062**
Sortino Ratio (MAR = 0.00)	0.060	0.085	0.037	0.064	**0.095**
Information Ratio	-	0.370	-0.130	0.198	**1.038**
R-Squared	-	0.630	**0.297**	0.327	0.869
Skewness	-0.485	**2.166**	-1.310	-1.074	-0.872
Excess Kurtosis	**5.283**	41.077	239.272	113.567	10.314
Maximum Drawdown	**0.230**	0.295	0.610	0.469	0.276
Historical VaR (0.95)	-0.018	-0.017	-0.017	-0.017	**-0.016**
Historical ES (0.95)	-0.028	**-0.026**	-0.034	-0.031	**-0.026**
Modified VaR (0.95)	-0.019	**-0.017**	**-0.017**	-0.018	**-0.017**
Modified ES (0.95)	-0.026	**-0.025**	-0.026	**-0.025**	**-0.025**
Average Monthly Turnover	-	0.125	0.133	1.262	**0.072**

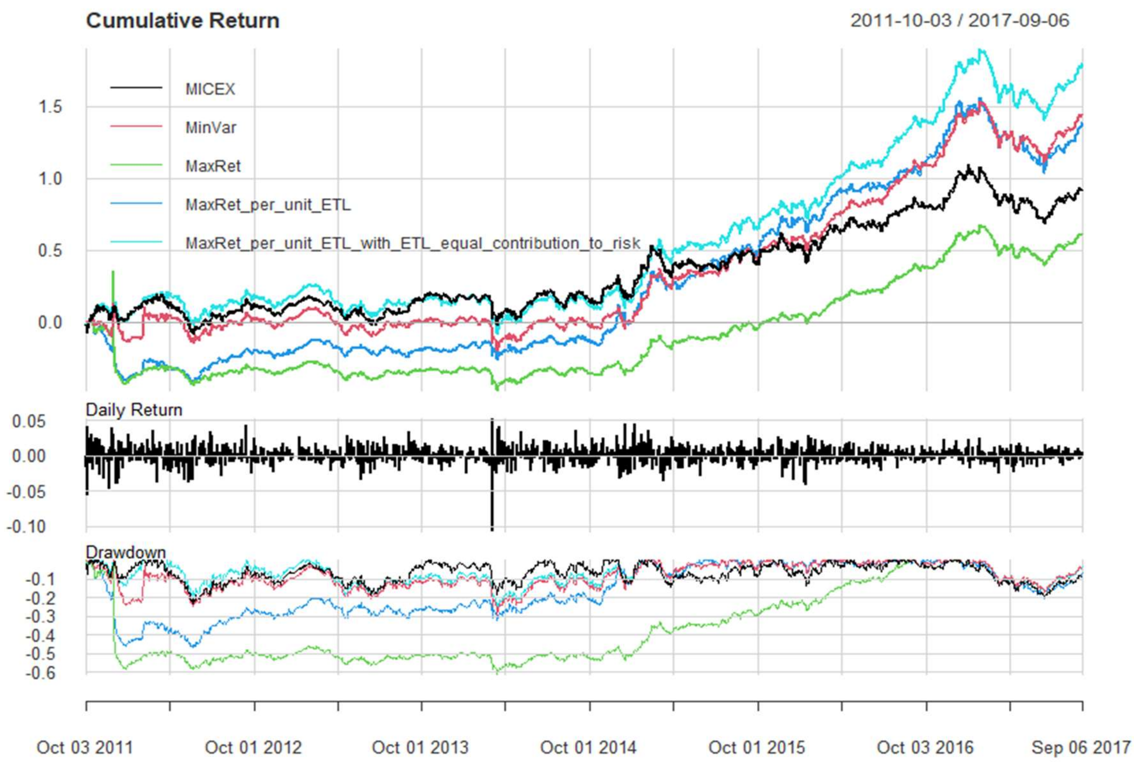

Figure 4.1 Performance summary of the classical mean-risk portfolios with a proportional transaction cost constraint and monthly rebalancing

In conclusion, this first application has shown that the long-term performance of an optimal mean-risk portfolio can be improved by adjusting the component weights to match the desired risk profile better. Here, each asset was required to contribute an equal amount of risk to the portfolio. Indeed, as the number of assets held in any portfolio increases, many possible portfolios will have similar mean return and standard deviation. Therefore, it is essential to be able to add information to the portfolio selection process. Consistent with Boudt et al. (2008, 100), we find that the modified Expected Shortfall and portfolio risk decomposition techniques help achieve such a task. Also, imposing bound constraints of the percentage ES contributions can be seen as a direct substitute for a risk diversification approach based on position limits (Boudt et al. 2013, 48). Finally, our analysis brings new perspectives to the earlier findings of Brandt (2005) and Martellini and Ziemann (2008, 13). They argue that global minimum variance portfolios typically achieve higher out-of-sample Sharpe ratios than other tangency portfolios in a mean-variance portfolio choice.

However, when utilizing the component decomposition of risk, one should note that for ES (and VaR), the capital allocation given by the portfolio weights may be very different from the risk exposures. Also, the estimated risk allocation depends on the risk measure used. A technical explanation for this follows from the definition in (3.2.7). In other words, asset i's percentage risk contribution is defined as the derivative of the risk measure with respect to the weight of that component multiplied by the component's weight in the portfolio and divided by the value of that risk measure. For example, in an equally-weighted portfolio, the expected contribution to the portfolio downside risk will be higher for assets with negatively skewed and thick-tailed returns than for assets with normal returns or positively skewed returns. (Boudt et al. 2008, 100.) Furthermore, for a portfolio that has the ERC constraint, the relative weights are inversely proportional to the marginal impact of the position on the portfolio ES, that is, $w_i/w_j = (\partial \mathrm{ES}/\partial w_j)/(\partial \mathrm{ES}/\partial w_i)$. Consequently, this allocation strategy produces portfolios that give higher weights to assets with a small marginal risk impact and down-weights the investments with a high marginal risk (aka 'hot spots' in Litterman 1996). In other words, it seeks positions for which a marginal decrease in weight leads to a large reduction in ES. (Boudt et al. 2013, 65.)

With respect to future research, there are at least a couple of aspects to consider. Firstly, note that here we have modeled only the proportional transaction costs using an optimization constraint. A possible modification to this constraint would be to assume that the rate of increase in transaction costs changes at certain threshold points. In other words, as the trading size increases, it should become increasingly more costly to trade because of the market impact of the trade. On the other hand, since it costs more to trade small-cap stocks than to trade large-cap stocks, one might wish to calculate the average trading volume of each stock and assume that the trading cost is inversely proportional to the trading volume. Alternatively, one could follow Fabozzi, Focardi, and Kolm (2006, 81–82), who define an optimal execution strategy that considers the variance of the total transaction costs. In contrast, Konno and Wijayanayake (1999) propose a branch and bound algorithm for solving a mean-absolute deviation portfolio optimization model, assuming that the cost function is concave (non-convex). For more discussion on different ways to incorporate transaction costs into standard portfolio allocation models and the impact of costs on portfolio optimization, see, e.g., Grinold and Kahn (2000, 450–452), Krokhmal, Palmquist and Uryasev (2001, 17), Scherer and Martin (2005, 98–105), Chincarini and Kim (2006, Chapter 10), Mansini, Ogryczak and Speranza (2015, 49–53, 59–61), as well as Pachamanova and Fabozzi (2016, 298–303).

Another issue that should be addressed in more detail in future research is the choice of rebalancing frequency. For instance, Tokat and Wicas (2007) identify the factors influencing a rebalancing strategy and present a conceptual framework for developing such strategies. Their findings indicate that determining an effective rebalancing strategy is a function of the time horizon and the characteristics of a portfolio's assets: their expected returns, volatility, and the correlation of their returns. In addition to this, an appropriate rebalancing strategy depends on return patterns over time and the market environment. (see also Liu & Loewenstein 2002; Zakamouline 2002; Pliska & Suzuki 2004; Jaconetti, Kinniry & Zilbering 2010.) Low et al. (2018, 279), in turn, observe that the degree of variance and the maximum positive or negative adjustment in portfolio weights required to achieve the desired investment objective tends to decrease as the size of the portfolio increases. Finally, Donohue and Yip (2003) consider optimal portfolio rebalancing with transaction costs and show that a strategy defining a no-trade region around target ratios can improve on the simple calendar- and volatility-based strategies.

Next, let us conduct another backtest of minimum-variance portfolio optimization using both the classical and robust estimators for moment estimation and analyze whether there is any difference in relative performance. The optimization is carried out by imposing the full investment and long-only constraints. Figure 4.2 depicts the excess returns of the backtest based upon the robust statistics compared to the classical estimator, and Table 4.2 lists the relevant summary statistics. The most preferred values are bolded.

As can be seen, the use of the S-estimator led to the least volatile portfolio allocations throughout the backtesting period. In other words, although this estimator produced robust portfolios with the lowest median (0.002), mean (0.004), third quartile (0.079), and maximum (2.215) excess returns, the difference between the maximum and minimum excess returns is the smallest (3.512). On the other hand, the distribution of excess returns has the lowest skewness of 1.425. The S-estimator is followed by the MM-estimator with the second smallest range of 5.719 and the highest skewness of 2.714. The Minimum Volume Ellipsoid (MVE) estimator, in turn, produced portfolio allocations that yielded excess returns of the second-highest skewness of 2.701. Furthermore, this is the estimator with the highest median (0.011) and mean (0.022) excess returns, as well as the highest maximum excess return of 4.920, which was generated during the ruble crisis in November-December 2014.

In contrast, the Minimum Covariance Determinant (MCD) estimator produced the most volatile portfolio allocations, as is illustrated by the most negative excess return peaks, for instance, at the end of 2012, as well as during the turmoil period at the end of 2014. Indeed, these robust portfolios have the lowest minimum (-2.066) and first quartile (-0.172) excess returns but also the highest third quartile excess return of 0.195, leading to the largest excess return range of 6.736. Finally, the Orthogonalized Gnanadesikan-Kettenring (OGK) and M-estimator produced robust portfolios that are between these extremes and fairly similar to each other, e.g., in terms of the excess return range (6.148 and 6.307, respectively). Both also have the same second-highest median excess return of 0.008.

In conclusion, the outcome of the minimum variance optimizations based on the robust estimators is more favorable than that of the classic estimator for the given measures. The median and average excess returns are positive for all robust portfolios, and they are skewed to the right. Of course, the skewness is affected by the large positive excess returns at the end of 2014 and at the beginning of 2015. However, positive skew results even if these data points are disregarded. In other words, the difference between the third quartile and the absolute value of the first quartile is greater than zero for all the other robust estimators, except for the S-estimator (-0.001). In this respect, the S-estimator is the least favorable compared to the other robust estimators. The MVE estimator produces the highest value of 0.032, and it is followed by the MCD estimator (0.023).

Our findings are in line with Pfaff (2016, 183–186), who demonstrates that even when the Gaussian copula with normally distributed margins is used to generate random asset returns, the classical covariance estimator may not generate a portfolio structure with the lowest risk on average. Moreover, the dispersion of the portfolio risk may neither be the smallest. However, with respect to the median risks, the differences become negligible as the sample size increases.

Table 4.2 Excess returns of the backtest based on the robust statistics versus the classical estimator

Robust estimator	*Minimum*	*1st quartile*	*Median*	*Mean*	*3rd quartile*	*Maximum*	*Skew*
M	-1.968	-0.160	0.008	0.020	0.181	4.339	2.037
MM	-1.709	-0.113	0.007	0.015	0.119	4.010	**2.714**
MVE	-1.760	-0.140	**0.011**	**0.022**	0.172	**4.920**	2.701
MCD	-2.066	-0.172	0.005	0.020	**0.195**	4.670	1.994
S	**-1.297**	**-0.080**	0.002	0.004	0.079	2.215	1.425
OGK	-1.813	-0.153	0.008	0.021	0.159	4.335	2.034

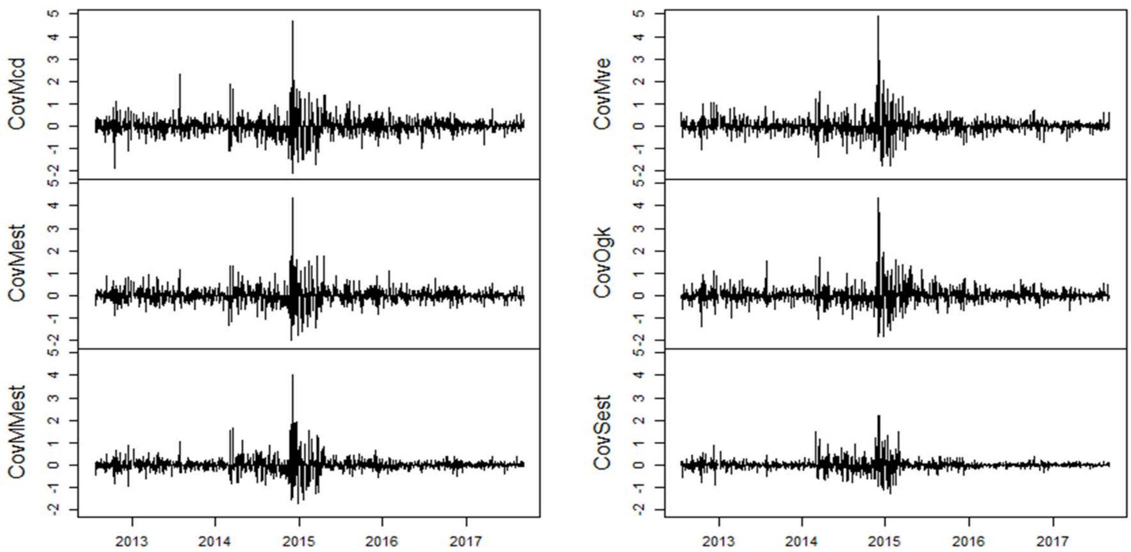

Figure 4.2 Relative performance of robust minimum variance portfolios

The applicability of robust estimators can be illustrated further. Utilizing the R packages *PortfolioAnalytics* (Peterson et al. 2018) and *MASS* (Ripley et al. 2021), we now write a custom moment function to estimate the variance-covariance matrix of the asset returns using the robust Minimum Volume Ellipsoid (MVE) method. Again, we construct monthly rebalanced portfolios where the optimization objective is to minimize portfolio variance. The robust estimates are compared with the sample estimates in two instances: In the first case, we apply the long-only and full investment constraints. In the second case, we allow short selling, but to avoid concentration risk, we add a box constraint for the asset weights so that the minimum weight of any asset must be greater than or equal to minus 25 percent, and the maximum weight of any asset must be less than or equal to 25 percent. If the weights for one or more periods do not sum up to one, we assume that the uninvested capital does not earn any interest. The size of the rolling window of data (i.e., the lookback window), as well as the length of the training period, are both set to 252 days. In order to evaluate whether these optimization models can outperform a naïve weighting scheme, we also show the performance of an equal-weighted benchmark strategy.

Furthermore, in both cases, we omit the use of the proportional transaction cost constraint. We do so because too binding constraints tend to restrict the Differential Evolution optimization algorithm (see Mullen et al. 2011) to such an extent that it resorts to using equal weights for each asset on consecutive rebalancing periods. Naturally, this kind of outcome makes it more challenging to see the differences in portfolio optimizations when using either a robust estimator or the classical sample estimator. Nevertheless, we report the average monthly turnover rates and treat them as a proxy for transaction costs. It is also worth mentioning that we experimented with a couple of alternative specifications. First, we tried more frequent (i.e., weekly) rebalancing. Second, instead of requiring that the portfolio weights must sum to one, we used the dollar neutral (or active) constraint (see Chalabi & Würtz 2015, 460; Garcia 2020; Ripley et al. 2021, 32–34 for more technical details). However, since those could not outperform the current setting, we do not show their results here.

The statistics used to evaluate the out-of-sample performance are reported in Table 4.3. Correspondingly, the evolution of the cumulative return and drawdown for each portfolio is illustrated in Figure 4.3. The first observation is that each of the four variants of the minimum variance strategy outperforms the equally-weighted portfolio in terms of the annualized Sharpe ratio, Sortino ratio, skewness, maximum drawdown, and modified Expected Shortfall (95%). When focusing on the differences between the long-only portfolios, we can see that the use of the robust estimator increases the annualized Sharpe ratio from 1.91 to 2.14, i.e., the increase in the annualized return is greater than the increase in the annualized standard deviation. Interestingly, the robust long-only portfolio even has a higher annualized Sharpe ratio than the classical minimum variance portfolio with short-selling allowed (2.09). This is because in the latter case, the only 2.1 percentage point increase in the annualized return is achieved by increasing the annualized standard deviation by 1.4 percentage points. Also, only when short-selling is forbidden are the minimum variance portfolios less risky than the equally-weighted benchmark in terms of the annualized standard deviation. Finally, when comparing the portfolios with short-selling allowed, we can see that in the case of the robust portfolio, the annualized return increases by five percentage points compared to the classical counterpart, but the annualized standard deviation decreases by 0.40 percentage points, leading to an increase in the annualized Sharpe ratio from 2.09 to 2.44.

Similar conclusions can be made by inspecting the Sortino ratio, i.e., the robust portfolio with short-selling allowed has the highest value of 0.20 and is followed by the robust long-only variant (0.19). These portfolios also have the highest information ratios of 1.61 and 1.16, respectively. In terms of the R-squared, the classical minimum variance portfolio with short-selling allowed is the least correlated with the benchmark portfolio (0.46), whereas the opposite is true for the classical long-only variant (0.60). Furthermore, the robust long-only portfolio is the only one with positively skewed returns (0.14), whereas the classical long-only portfolio has a return distribution with the smallest kurtosis of 6.59. On the other hand, when short-selling is allowed, the switch from the classical estimator to the robust estimator decreases the kurtosis from 15.54 to 8.21.

In terms of the downside risk measures, the robust long-only minimum-variance portfolio outperforms both the competing variants and the benchmark portfolio. More precisely, it has the least severe maximum drawdown of 0.15, as well as the smallest historical Expected Shortfall (-0.02), modified Value-at-Risk (-0.01), and modified ES (-0.02). Correspondingly, the robust portfolio with short-selling allowed has 4.3 percentage points lower maximum drawdown than its classical counterpart, although its modified Expected Shortfall is slightly higher (-0.04 and -0.03, respectively). Finally, the benchmark portfolio has the lowest average monthly turnover of 0.06, and it is followed, surprisingly, by the classical long-only minimum-variance portfolio (0.98). Indeed, both the robust long-only portfolio (1.09) and the robust portfolio with short-selling allowed (2.27) have higher turnover rates than their classical counterparts.

Therefore, two conclusions can be made from this robustification exercise. Firstly, the utilization of the robust estimates of the variance-covariance matrix (as opposed to the ordinary sample estimates) produces portfolios that outperform the classical minimum-variance portfolios in terms of risk-adjusted returns. Moreover, as can be seen from Figure 4.3, the cumulative returns of the robust long-only and the classical portfolio with short-selling allowed are fairly close to each other (3.35 and 3.70, respectively), meaning again that shorting is not the only way to improve the returns produced by the minimum variance strategy. Secondly, if the higher average turnover values are excluded, it can be argued that the downside risk profile of a robust portfolio is more attractive than that of its classical counterpart, especially when short-selling is not allowed.

Table 4.3 Performance metrics of the minimum variance portfolios with classical sample or robust MVE estimators versus an equally-weighted portfolio

	Equal-weighted	*Long-only Min var (sample)*	*Long-only Min var (robust)*	*Min var (sample)*	*Min var (robust)*
Cumulative Return	1.534	2.590	3.349	3.703	**4.655**
Annualized Return	0.201	0.286	0.336	0.357	**0.407**
Annualized Std Dev	0.165	**0.150**	0.157	0.171	0.167
Annualized Sharpe (Rf=0.00)	1.220	1.914	2.136	2.086	**2.441**
Sortino Ratio (MAR = 0.00)	0.107	0.163	0.193	0.178	**0.202**
Information Ratio	-	0.804	1.159	1.158	**1.608**
R-Squared	-	0.602	0.547	**0.462**	0.492
Skewness	-1.118	-0.612	**0.138**	-0.670	-0.667
Excess Kurtosis	15.130	**6.588**	7.175	15.544	8.211
Maximum Drawdown	0.276	0.148	**0.145**	0.205	0.162
Historical VaR (0.95)	-0.015	**-0.013**	**-0.013**	-0.015	-0.015
Historical ES (0.95)	-0.023	-0.022	**-0.020**	-0.022	-0.023
Modified VaR (0.95)	-0.016	-0.015	**-0.013**	-0.015	-0.016
Modified ES (0.95)	-0.043	-0.031	**-0.018**	-0.029	-0.036
Average Monthly Turnover	**0.063**	0.980	1.090	2.079	2.265

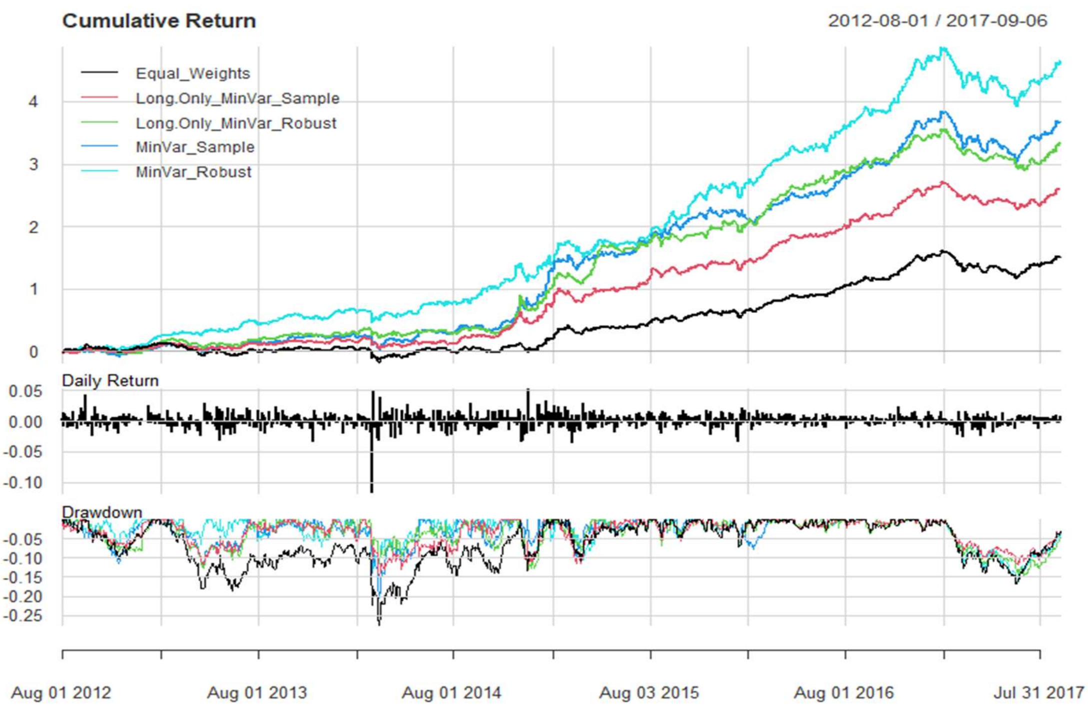

Figure 4.3 Performance summary of the minimum variance portfolios with classical sample or robust MVE estimators versus an equally-weighted portfolio

With respect to future research, it is worth mentioning that we have omitted the implementation of at least two well-known robust estimators of the covariance matrix, namely the shrinkage estimator and the projection-based Stahel–Donoho estimator (SDE). In brief, the benefit of the shrinkage estimate is that it is always positive definite, well-conditioned, and efficient. However, although the resulting regularized estimator may outperform the maximum likelihood estimator for small samples, the shrinkage intensity will reduce to zero for large samples. This means that the shrinkage estimator will be identical to the empirical estimator. (see Würtz et al. 2009, 289–292.) Nevertheless, the interested reader wishing to implement such an approach is referred, e.g., to Schäfer and Strimmer (2005) and Opgen-Rhein and Strimmer (2007). More technical information on the efficient estimation of a James-Stein-type shrinkage estimator for the covariance matrix, with separate shrinkage for variances and correlations, can also be found from the vignette of an R package called *corpcor* (Schäfer et al. 2017). Similarly, an R package called *tawny* (Lee & Rowe 2016) allows one to carry out portfolio optimizations based on shrinkage estimators. Considering the SDE estimator, one should note that although it possesses good statistical properties, it is extremely slow when the number of variables is large. Hence, it is better to use the MCD or OGK estimator for larger data sets (Todorov 2020, 60). For more information on the SDE method, see, e.g., Stahel (1987) and Donoho and Gasko (1992).

The final part of this subchapter considers the backtest results produced by the robust portfolio optimization approach with elliptical uncertainty of $\boldsymbol{\mu}$. Based on the computation method proposed by Pfaff (2016), Figure 4.4 displays the efficient frontier with equivalence points (values in percentages). As in general, the x-axis represents the risk measured by the standard deviation ($sigma$, σ), whereas the y-axis represents the expected return (mu, μ). The efficient frontier is drawn as a line for the calculated points as per the classical mean-variance (MV) allocations. The efficient points regarding the MV and robust counterpart (RC) allocations are superimposed. These points are depicted as blue and red dots, respectively. Finally, drawn in are the equivalent MV and robust counterpart points. This is done such that first, the RC allocation with the trade-off parameter (i.e., risk aversion) $\theta = 0.7$ is shown as a green square on the efficient frontier. Then, the equivalent classical MV portfolio is portrayed as an orange square on the efficient frontier.

We can see that the selected robust optimization solutions fall short of the MV solutions, although they are part of the MV efficient frontier. This becomes evident when we compare the *equivalent solutions* of the MV and robust optimizations. Indeed, the solution shown as a point to the west is the equivalent MV solution to the neighboring robust counterpart, which is located to the southwest of it. Such a conclusion is consistent with Schöttle (2007, 125–126) and Pfaff (2016, 194).

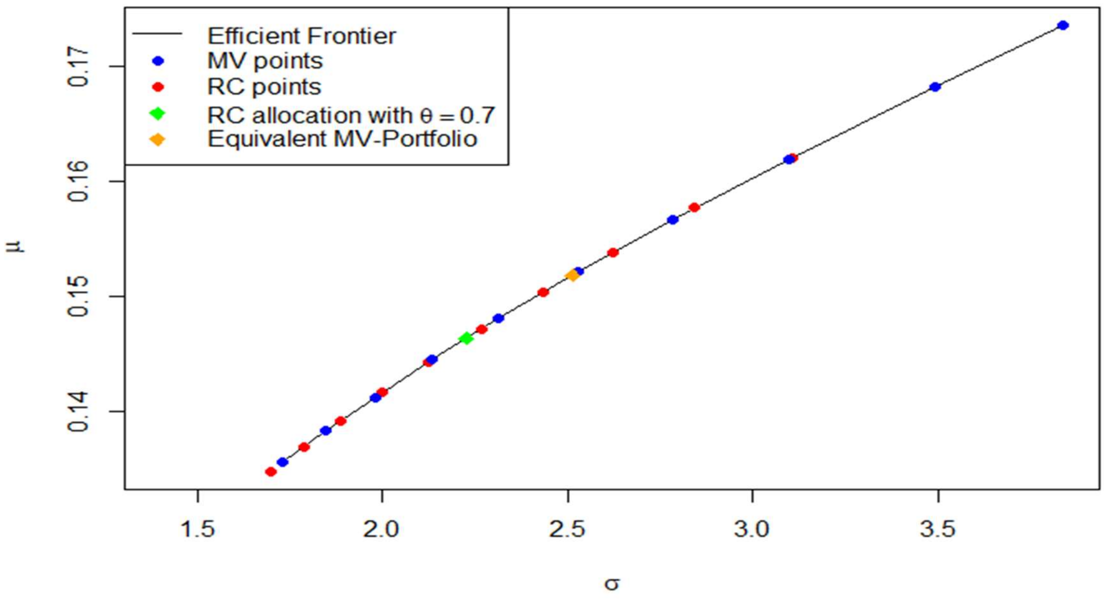

Figure 4.4 Efficient frontier of classical and robust MV portfolios

Indeed, in this particular setting, we have discovered that the robust efficient frontier is identical to the classical efficient frontier, but it is shortened in respect to the risk axis. Hence, it does not attain portfolios with as high risk as the maximum return portfolio in the classical framework. Technically speaking, robust efficient frontier coincides with the classical efficient frontier up to the trade-off parameter (risk level) $\lambda = 1/\left(1 + \frac{\delta}{\sqrt{T}}\right)$ of the classical frontier. For the trade-off parameter $\theta = 1$ at the right end of the robust efficient frontier, the maximum value for λ is reached. (see Schöttle 2007, 145–146.)

Correspondingly, to show that the classical and robust efficient frontier coincide for the entire length of the robust efficient frontier, it suffices to prove that the optimization problems to discover the explicit points are equivalent. Again, using the tracing parameter defined in (3.2.27), the classical problem can be reformulated such that it is equivalent to the robust formulation. Thus, the only difference between these two optimization problems is the fraction $1/(1 + \theta\frac{\delta}{\sqrt{T}})$, which is just a constant. (see Schöttle 2007, 125–127 for the formal proof.)

Therefore, a few additional remarks can be made. Firstly, if we study the special point $\theta = 0$, that is, the (robust) minimum variance portfolio, we can be seen that the corresponding λ is also zero and both programs reduce to $\min_{\boldsymbol{\omega}\in\Omega}\sqrt{\boldsymbol{\omega}'\hat{\Sigma}\boldsymbol{\omega}}$. The equivalence of the classical and robust minimum variance portfolio is expected because the uncertainty of the covariance matrix was not explicitly considered, and the minimum variance portfolio does not depend on the uncertain parameter $\boldsymbol{\mu}$. Secondly, when applying the RC approach in the general case with $\theta \in [0, 1]$, we attain the expression $\lambda\delta\sqrt{\boldsymbol{\omega}'\hat{\Sigma}\boldsymbol{\omega}}$. It operates as a type of regularization and can be understood as estimation risk. According to Schöttle (2007, 127), this additional expression penalizes the investment in stocks with high volatility. In other words, it increases the impact of the portfolio risk compared to the expected return. This kind of robustification implies a change from the trade-off between risk and return towards the less risky portfolios.

For a particular investor, the implication is as follows: Suppose an investor has a specific risk aversion parameter λ which determines her personal trade-off between return and risk. By robustifying the portfolio optimization problem with a confidence ellipsoid, the investor changes her position on the efficient frontier. In other words, she becomes more conservative, and hence her optimal robust portfolio moves towards the minimum variance portfolio (MVP). It can be further shown that since the robust optimal portfolio is a bit closer to the MVP, the investment in the more secure assets is typically higher. (see Schöttle 2007, 128.)

With respect to future research, it is worth highlighting that here (as is common in practical problems), we have only determined an uncertainty set for the vector of expected returns. This decision is justified because the covariance matrix is not as volatile and does not as crucially affect the optimal solution as the return estimate. Nevertheless, it would be possible to consider the covariance matrix as an uncertainty parameter as well. Without going into details, such an extension requires the definition of a joint confidence ellipsoid around a point estimate for the pair $(\boldsymbol{\mu}, \Sigma)$, and thus one needs to determine the distribution of the point estimate $(\hat{\boldsymbol{\mu}}, \hat{\Sigma})$. As shown in Schöttle (2007, 129–148), robustifying both parameters leads to a more conservative portfolio allocation than robustification of $\boldsymbol{\mu}$ only, but this distinction is usually unrecognizably small in practice.

Alternatively, one could create an ellipsoidal uncertainty set for the unknown return vector $\boldsymbol{\mu}$ by utilizing multiple robust-based estimators (such as $\hat{\boldsymbol{\mu}}^{M}$, $\hat{\boldsymbol{\mu}}^{MM}$, $\hat{\boldsymbol{\mu}}^{MVE}$, $\hat{\boldsymbol{\mu}}^{MCD}$, $\hat{\boldsymbol{\mu}}^{S}$, $\hat{\boldsymbol{\mu}}^{OGK}$) or their combination for this parameter. It can be further shown that in the case of various estimators, the robust optimization problem penalizes investment in assets where the considered statistical estimators generate different values (i.e., lie further apart from each other). Moreover, using an uncertainty set based on estimators leads to a robust efficient frontier that is shorter than the classical one, meaning that when the investor's expected return requirement increases above a certain threshold, she can utilize the classical optimization to obtain the same efficient allocation as in the case of robust optimization but with lower risk (measured by standard deviation). (see Schöttle 2007, 103–106, 148–158; Pfaff 2016, 170 for more discussions.)

Let us now investigate the dynamic behavior of classical (termed Markowitz) and robust mean-variance portfolio optimizations and see how they perform in comparison to the competing investment strategies. As explained in Subchapter 3.2, we first computed the initial asset allocations for each of the five strategies. Based on these initial portfolio weights, we found that the Maximum Sharpe Ratio (MSR) and Markowitz optimizations led to the most concentrated portfolio solutions (not depicted here). More precisely, the maximum allowed weight of 25 percent was reached for three stocks in the case of the

MSR portfolio. These assets were PhosAgro, Surgutneftegas, pref., and Transneft, pref. In the case of the Markowitz portfolio, Alrosa also achieved the weight limit in addition to the above three assets. The initial weight allocations were more diversified in the portfolios based on the inverse variance and robust optimization strategies, where no single asset received a weight of over ten percent.

The backtest results are summarized in Table 4.4. First, recall that the return figures include the transaction fees of 100 basis points (one percent) and that the risk-free interest rate is set to one percent per annum. Second, in addition to the performance and risk metrics introduced earlier, we use MATLAB® to compute the *average turnover* (the average per-time-step portfolio turnover), *max turnover* (the maximum portfolio turnover in a single rebalance), *average buy cost* (the average per-time-step transaction costs the portfolio incurred for asset purchases), and *average sell cost* (the average per-time-step transaction costs the portfolio incurred for asset sales). The most appealing values have been bolded. Figure 4.5. shows the cumulation of portfolio returns, the daily returns of the equal-weighted benchmark strategy, and the underwater chart for drawdowns. Finally, Figure 4.6 visualizes the time-varying Value-at-Risk estimates.

It can be seen that although the MSR and Markowitz optimization strategy both achieve the highest annualized return of 0.32, the former has about six percentage points lower annualized standard deviation than the latter (0.17 and 0.23, respectively). Hence, the MSR portfolio has (perhaps unsurprisingly) the highest annualized Sharpe ratio of 1.83. This tangency strategy is then followed by the robust optimization approach (1.40) and its Markowitzian counterpart (1.34). Indeed, in line with the previous application, there is a risk-return tradeoff when choosing between the classical and robust mean-variance optimizations. More specifically, the lower annualized return of 0.23 in the case of the robust portfolio is compensated by the lowest annualized standard deviation of 0.15. Finally, in terms of risk-adjusted returns, even the equal-weighted portfolio (1.09) manages to outperform the MICEX index (0.64). The inverse variance strategy has the same annualized return as the naïve allocation strategy (0.19) but roughly one percentage point lower annualized standard deviation (0.16 and 0.17, respectively).

Correspondingly, the MSR strategy produces the highest Sortino ratio (0.16) and is followed by the robust optimization approach (0.13) and its Markowitzian counterpart (0.12). On the other hand, using the MICEX index as a benchmark, the robust portfolio has a higher information ratio than the MSR portfolio (1.49 and 1.33, respectively). This implies that the former strategy achieves a higher return in excess of the benchmark, given the risk taken. In terms of the R-squared, the Markowitz optimization approach (0.31) is the least correlated with the MICEX index, whereas the opposite is true for the inverse variance strategy (0.90). Interestingly, both the robust optimization and equal-weighted strategy have a high and roughly similar coefficient of determination (0.89 and 0.86, respectively), whereas the R-squared is only 0.44 in the case of the MSR portfolio. Thus, only 44 percent of the MSR portfolio's performance can be attributed to the performance of the MICEX index.

Based on the distribution characteristics, the returns of the MICEX index are the least negatively skewed (-0.49), and they have the lowest kurtosis of 6.91. It is followed by the MSR portfolio with a skewness of -0.64 and a kurtosis of 7.40. When comparing the classical and robust mean-variance optimization approaches, it can be seen that the latter manages to increase the skewness from -1.13 to -0.89 and lower the kurtosis from 12.04 to 10.27. Furthermore, with respect to the downside risk measures, the Markowitz portfolio has the worst maximum drawdown of 0.33, whereas the robust counterpart has the smallest value of 0.18. The robust portfolio also has the least severe historical Expected Shortfall of -0.021, whereas the opposite is true for the Markowitz portfolio (-0.032). In terms of the historical VaR (-0.014) and modified VaR (-0.015), both the robust optimization and inverse variance strategy are the least risky. Finally, considering the modified Expected Shortfall, both the MICEX index and MSR portfolio produce the least negative value of -0.036, followed by the robust portfolio (-0.037) with only one-tenth of a percent difference.

By inspecting the drawdown chart in Figure 4.5, it can be observed again that the most drastic losses occurred around March 2014. During that time, the Markowitz portfolio was the only one exceeding the drawdown level of -0.30. Although the simultaneity of loss periods is visible, it can also be noticed that during the ruble crash in fall 2014, the drawdown of the Markowitz portfolio reaches the -0.30 threshold for the second time. In contrast, the MSR strategy stays below the -0.20 level, whereas the inverse variance and robust optimization strategies are closer to the -0.05 level. Moreover, during the first half of 2017, the Markowitz portfolio solution was once again the riskiest as the drawdowns peaked twice, reaching the level of -0.25. At the same time, the MSR portfolio was the least volatile as its drawdowns did not drop below the -0.10 threshold.

According to the estimated VaR values at the 95 percent confidence level in Figure 4.6, the classical mean-variance optimization approach consistently produces the highest estimates throughout the backtesting period. For instance, in March 2014, it generates the highest peak of 0.065, whereas the other strategies stay below the 0.05 level, excluding the equal-weighted portfolio (0.052). Similarly, in September 2014, the Markowitz portfolio exceeds the 0.05 threshold, unlike the other strategies staying below the 0.02 level, excluding the MSR strategy (0.026). Thus, in terms of the average VaR estimates, the Markowitz portfolio can be considered the riskiest (0.025), and it is followed by the MSR portfolio (0.017). The robust optimization and the inverse variance strategy both have the lowest average value of 0.015.

Regarding the average turnover, the lowest value of 0.002 is produced by the equal-weighted, inverse variance, and robust optimization strategies. In contrast, the Markowitz optimization strategy has the highest average turnover of 0.012. Unsurprisingly, the equal-weighted portfolio also has the smallest maximum turnover of 0.083, whereas the Markowitz portfolio has the greatest value of 0.538. Similarly, the equal-weighted strategy has the lowest average buy (and sell) cost of 209.15, and the opposite is true for the Markowitz portfolio (2341.03). The MSR portfolio, in turn, produces the second-highest average and maximum turnover rates of 0.01 and 0.49, respectively, as well as the second-highest average buy (and sell) cost of 1988.72.

Again, when comparing the classical and robust optimization approaches, it can be seen that robustification decreases the turnover and transaction costs significantly. More specifically, the average and maximum turnover decrease by roughly 83 and 81 percent, respectively, whereas the average buy (and sell) cost decreases by about 84 percent. Hence, in line with the findings of Schöttle (2007, 157–158), as the turnover of the robust portfolio is smaller than the turnover of the Markowitz portfolio, the respective costs are more limited, resulting in a smaller performance reduction compared to the classical mean-variance approach. Finally, the robust optimization approach creates more stable portfolios that reasonably meet their expectations under the Russian market's volatile conditions.

In conclusion, we have now seen that the strategy maximizing the Sharpe ratio produces the highest risk-adjusted returns, and it is also the least risky in terms of the modified Expected Shortfall at the 95 percent confidence level. The robust mean-variance optimization approach, in turn, improves the classical counterpart by decreasing the risk measured by the annualized standard deviation, leading to the second-highest risk-adjusted returns on average. The robust portfolio is also less risky than the Markowitz portfolio regarding the downside risk measures given. Therefore, it can be argued that robustification adds value in asset management and is an especially suitable investment solution for conservative and risk-averse investors.

A natural way to further explore this topic would then be to consider robust portfolio optimization models for various reward–risk ratios, such as Omega ratio, semi-mean absolute deviation (SMAD) ratio, and weighted stable tail adjusted return ratio (WSTARR). The interested reader is referred, e.g., to Sehgal and Mehra (2021), who address the uncertainty in returns on assets by varying them in symmetric bounded intervals. In line with our findings, they find that the robust reward–risk ratio models outperform their conventional counterparts in terms of the standard deviation and downside risk measures.

Table 4.4 Performance results from the backtest of the robust optimization and four competing strategies versus the MICEX index

	MICEX index	Equal-weighted (TC=100 bps)	Max Sharpe ratio (TC=100 bps)	Inverse variance (TC=100 bps)	Markowitz optimization (TC=100 bps)	Robust optimization (TC=100 bps)
Cumulative Return	0.848	1.439	3.118	1.431	**3.124**	1.815
Annualized Return	0.128	0.191	**0.320**	0.190	**0.320**	0.225
Annualized Std Dev	0.183	0.165	0.168	0.155	0.230	**0.152**
Annualize Sharpe (Rf=0.01 p.a.)	0.639	1.087	**1.827**	1.150	1.336	1.403
Sortino Ratio (MAR = 0.00)	0.068	0.102	**0.163**	0.108	0.118	0.128
Information Ratio	-	0.909	1.325	1.035	0.965	**1.488**
R-Squared	-	0.856	0.438	0.902	**0.306**	0.885
Skewness	**-0.487**	-1.086	-0.640	-0.976	-1.134	-0.892
Excess Kurtosis	**6.911**	14.228	7.404	12.266	12.041	10.266
Maximum Drawdown	0.196	0.280	0.252	0.208	0.332	**0.183**
Historical VaR (0.95)	-0.017	-0.015	-0.015	**-0.014**	-0.020	**-0.014**
Historical ES (0.95)	-0.025	-0.024	-0.023	-0.022	-0.032	**-0.021**
Modified VaR (0.95)	-0.018	-0.016	-0.017	**-0.015**	-0.023	**-0.015**
Modified ES (0.95)	**-0.036**	-0.043	**-0.036**	-0.039	-0.062	-0.037
Average Turnover	-	**0.002**	0.010	**0.002**	0.012	**0.002**
Max Turnover	-	**0.083**	0.487	0.130	0.538	0.101
Average Buy Cost	-	**209.155**	1988.724	325.864	2341.027	381.329
Average Sell Cost	-	**209.155**	1988.724	325.864	2341.027	381.329

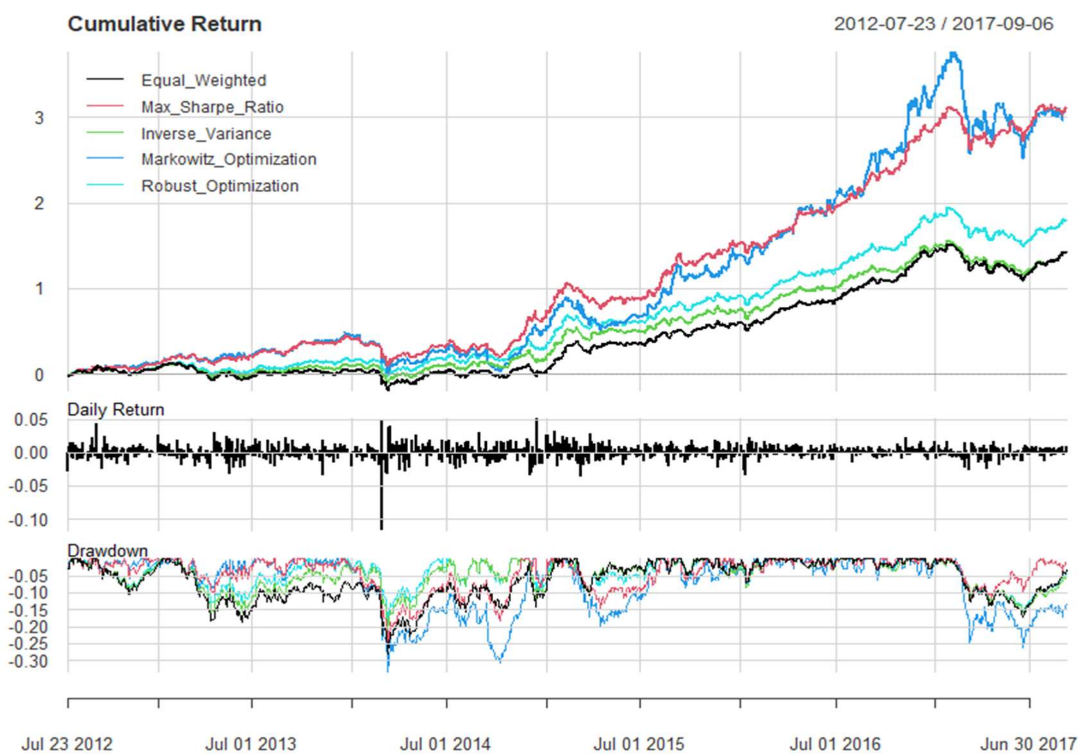

Figure 4.5 Performance summary of the robust optimization and four competing strategies over the five years

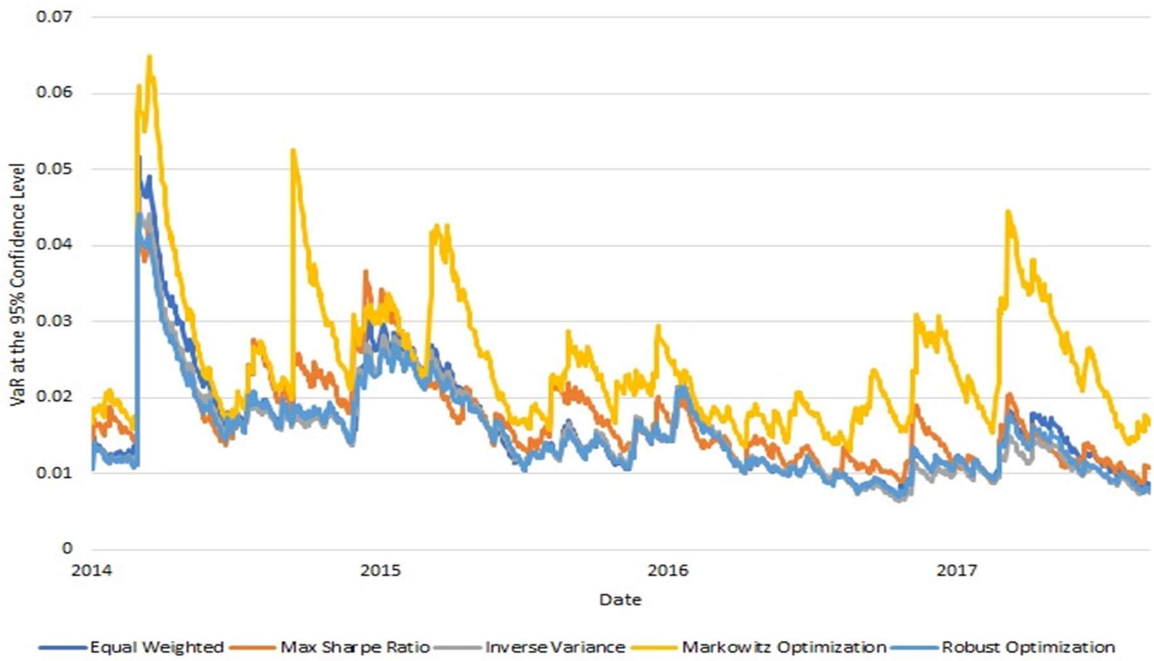

Figure 4.6 Value-at-Risk estimation for five competing strategies using the EWMA method

With respect to future research, it is also worth mentioning that the errors of the estimates for the expected returns and the variance-covariance matrix could have been quantified heuristically beforehand by means of Monte Carlo simulations. This is known as portfolio resampling (see Michaud 1998, as well as Scherer 2002; 2010 for a critical analysis). By utilizing the 'resampled efficiency' concept, randomized efficient portfolios can be retrieved by averaging the weights over K random sample pairs that are used to compute m points on the respective efficient frontiers. However, such an approach should not be considered as a universal remedy for coping with estimation errors. Indeed, the major disadvantage here is the propagation of errors: The initial values $(\hat{\mu}_0, \hat{\Sigma}_0)$ from which the random samples are generated are estimates themselves (for the theoretical moments $\boldsymbol{\mu}$ and Σ) and hence contaminated by estimation errors. Therefore, the initial errors are replicated by applying Monte Carlo methods. Furthermore, as shown in Scherer (2010, 133–136), the above procedure can yield unintuitive results in the case of constrained portfolio optimizations. Finally, Pfaff (2016, 51) points out that there is often a model error in addition to estimation errors due to the non-normality and non-stationarity features of return processes (see also Broadie 1993; Michaud & Michaud 2008, Chapter 6 for more discussion).

Our final discussion relates to the simple techniques one can utilize to improve further the performance of the strategies presented in this subchapter. These include the requirement of market neutrality and weight smoothing. Considering the first way, it is clear that the market itself constitutes a risk factor. If the assets in the portfolio are positively correlated with the market, then the portfolio would follow not only market growth but also market downturns. Of course, one way for the investors to avoid the latter situations is to construct portfolios that are uncorrelated with the market (i.e., market-neutral). In order to be market-uncorrelated, the portfolio must have zero *beta*, that is, $\beta_p = \sum_{i=1}^{n} \beta_i x_i = 0$, where $x_1, \ldots, x_n$ denote the proportions (weights) in which the total portfolio capital is distributed among n assets. Moreover, β_i are market betas of individual assets, $\beta_i = \mathrm{Cov}(r_i, r_M)/\mathrm{Var}(r_M)$, where r_M stands for the market rate of return, and r_i is the (random) rate of return of asset i. Using historical data, assets' betas, β_i, can be estimated such that $\beta_i = \left(\sum_{j=1}^{J}(r_{M,j} - \bar{r}_M)^2\right)^{-1} \sum_{j=1}^{J}(r_{i,j} - \bar{r}_i)(r_{M,j} - \bar{r}_M)$, where J is the number of historical observations, and $\bar{r}$ denotes the sample average, $\bar{r} = J^{-1} \sum r_j$. (see Krokhmal, Uryasev & Zrazhevsky 2005, 617–619.)

In order to provide some preliminary insights into the portfolio performance when imposing the market-neutrality (zero-beta) constraint, we solved the maximum Sharpe ratio (MSR) optimization problem with and without this constraint. As in the previous case, the training period and the rolling window were set to 252 days, and the portfolios were rebalanced monthly. For the sake of simplicity, we added only the full investment and long-only constraints and omitted the considerations about transaction costs and maximum weight limits. As a proxy for market returns, r_M, we used the historical returns of the MICEX index and set the risk-free interest rate to zero percent. Then, in line with Bennett (2014, 19), we added a constraint requiring that the portfolio beta is between -0.25 and 0.25, i.e., $-0.25 \leq \sum_{i=1}^{n} \beta_i x_i \leq 0.25$. Hence, this constraint ensures that the portfolio correlation with the market is bounded and close to zero.

Table 4.5 shows the performance and risk metrics of the MSR portfolio with and without the market-neutrality (zero-beta) constraint versus the MICEX index. It also reports the returns per calendar year from 2012 to 2017. The most desirable values are bolded. It can be seen that as the R-squared decreases from 0.51 to 0.05, the annualized return increases by about 14 percentage points from 0.31 to 0.45. At the same time, however, the annualized standard deviation increases by roughly 16 percentage points from 0.16 to 0.32, meaning that the annualized Sharpe ratio decreases by about 55 percentage points from 2.00 to 1.45. Clearly, the market-neutral MSR strategy is a lot riskier than the basic version without the zero-beta constraint. The same can also be concluded by inspecting the MSR portfolio's and its constrained counterpart's Sortino ratios (0.17 and 0.16, respectively) and information ratios (1.40 and 0.99, respectively), which are more favorable for the former approach.

On the other hand, despite the higher kurtosis, the returns of the market-neutral portfolio are positively skewed (1.93), as opposed to the standard MSR portfolio (-0.63). Moreover, imposing the market-neutrality constraint decreases the modified Expected

Shortfall by almost two percentage points from -0.031 to -0.013 but increases the historical ES by nearly two percentage points from -0.021 to -0.038. Although the difference between these values is not exceedingly large, they do lead to somewhat conflicting conclusions about which portfolio has greater downside risk. Nevertheless, the same cannot be argued when inspecting the maximum drawdowns. Indeed, here we can notice that the MSR portfolio with the market-neutrality constraint suffered an extremely severe maximum drawdown of 0.55 (in December 2014), which is 41 percentage points larger than the maximum drawdown of the standard MSR portfolio (0.14). Finally, since the MSR portfolio is significantly more diversified, it has about 54 percentage points greater average monthly turnover than its market-neutral counterpart (1.00 and 0.46, respectively).

Based on the annual returns, we can see that the market-neutral MSR strategy produced an inconsistent and unstable return stream throughout the backtest period. Specifically, the portfolio returns were negative in 2012 and 2013 (-0.01 and -0.08, respectively), meaning that the strategy underperformed both its standard variant and the MICEX index. However, the situation changed dramatically in the following years, when the annual returns of the market-neutral portfolio peaked, reaching the value of 0.59 in 2014 and 1.72 in 2015. Indeed, during the year of crises in 2014, the market-neutral MSR strategy outperformed its counterpart by about 27 percentage points, and, in 2015, the outperformance was even more remarkable–roughly 106 percentage points. Towards the end of the backtesting period, however, the difference in annual returns narrowed. In 2016, the returns decreased to 0.58, and thus the outperformance was only about 17 percentage points compared to the standard MSR portfolio. In 2017, the difference in annual returns was only about six percentage points.

Again, it is worth emphasizing that the market-neutral MSR strategy could not outperform its standard counterpart in terms of risk-adjusted returns, except during 2015. This is due to the fact that the market-neutral approach led to concentrated portfolio solutions where individual assets received excessive weights, thus causing the return stream to be more volatile. For instance, in 2015, the market-neutral strategy invested on average nearly 80 percent on Polyus. Hence, our observations contrast with Krokhmal et al. (2005, 628), who find that imposing the market-neutrality constraint adds to the stability of the portfolio's return and reduces portfolio drawdowns.

In order to save space, we leave more detailed investigations on the effects of imposing the 'low-beta' requirement (or other more complex tail-dependence constraints) on the asset allocation algorithms presented in this subchapter for future research. The interested reader is referred, e.g., to Pfaff (2016, 204–217), who uses a so-called (lower) Tail Dependence Coefficient (TDC) to construct minimum tail-dependent portfolios against a benchmark. In brief, the lower TDC can be utilized between a broader market aggregate and its constituents so that the selected stocks entering a (long-only) portfolio are solely least concordant only to negative market returns. Thus, the aim is to shelter an investor from simultaneously incurring index losses and the stocks held in the portfolio. Of course, this kind of lower-tail-dependence strategy contrasts with the low-beta portfolio because a low-beta strategy aims to select stocks that co-move less proportionally than the market in absolute terms. Unfortunately, the selection based on a low-beta strategy may deter stocks that are characterized by a high value of the upper tail dependence coefficient with the market. Therefore, such a portfolio allocation may miss on the upside.

Table 4.5 Performance metrics of the maximum Sharpe ratio portfolio with and without the market-neutrality (zero-beta) constraint versus the MICEX index

	MICEX index	*Max Sharpe ratio*	*Max Sharpe ratio with the market-neutrality constraint*
Cumulative Return	0.859	2.967	**5.706**
Annualized Return	0.130	0.312	**0.455**
Annualized Std Dev	0.182	**0.156**	0.315
Annualized Sharpe (Rf=0.00)	0.713	**1.999**	1.445
Sortino Ratio (MAR = 0.00)	0.070	**0.171**	0.155
Information Ratio	-	**1.403**	0.992
R-Squared	-	0.513	**0.047**
Skewness	-0.483	-0.629	**1.929**
Excess Kurtosis	**7.024**	12.236	13.886
Maximum Drawdown	0.196	**0.139**	0.548
Historical VaR (0.95)	-0.017	**-0.013**	-0.023
Historical ES (0.95)	-0.025	**-0.021**	-0.038
Modified VaR (0.95)	-0.018	-0.014	**-0.013**
Modified ES (0.95)	-0.036	-0.031	**-0.013**
Average Monthly Turnover	-	0.997	**0.458**
Annual returns:			
2012	0.078	**0.084**	-0.007
2013	0.063	**0.172**	-0.082
2014	-0.018	0.318	**0.589**
2015	0.323	0.652	**1.716**
2016	0.328	0.406	**0.579**
2017	-0.059	0.020	**0.079**

As the above example demonstrated, the weights of mean-variance (and more general mean-risk) portfolio optimizations tend to be characterized by wide spans and erratic behavior over time. In other words, such portfolio optimizations suffer more severely from estimation errors than the minimum variance ones because the weight vectors of the latter do not depend on the expected returns of the assets. This finding has been well-established in several empirical studies (see, e.g., Merton 1980; Chopra & Ziemba 1993; Jagannathan & Ma 2003; Ledoit & Wolf 2003; Kempf & Memmel 2006; DeMiguel et al. 2009; Frahm 2010; Lejeune 2011). Of course, as has been done earlier in this subchapter, one can impose restrictions on the weights in order to ameliorate the extreme asset allocations and significant fluctuations beforehand. For instance, Frost and Savarino (1988) find that restrictions such as long-only and bound constraints improve out-of-sample performance, whereas Eichhorn, Gupta, and Stubbs (1998) find that they are related to reduced portfolio risk. According to Pfaff (2016, 51–52), both of these empirical findings can be traced back to a smaller implied estimation error if restrictions are imposed on the weights.

Instead of using arbitrary box and group constraints, Würtz et al. (2009, 351–379) and Würtz et al. (2015b, 6–7) suggest utilizing a simple Exponential Moving Average (EMA) smoother (see, e.g., Klinker 2011) with different decay lengths in order to achieve better stability properties. We conducted a preliminary investigation of this approach by smoothing the weight changes of a tangency strategy, which is defined so that it invests in a portfolio that has the highest Sharpe ratio, and if such a portfolio does not exist, the minimum variance portfolio is taken instead. For the sake of simplicity, the sample covariance was used as the risk measure. The EMA smoother then applied either a single or double EMA filter to a vector of portfolio weights with the decay parameter (*lambda*, λ) controlling the amount of smoothing. By utilizing a rolling portfolio analysis, the backtest was conducted for the end-of-month windows with a fixed horizon of 12 months, shifted monthly.

Table 4.6 shows the out-of-sample performance results for the three variants of the long-only tangency strategy versus the MICEX index (the most attractive values are bolded). In the first case, single smoothing was applied, and λ was set to one month, meaning that the optimal weights were given the most freedom to fluctuate. In the second case, time decay was increased to three months. As argued in Würtz et al. (2009, 382–383), setting lambda to three months is a reasonable choice if there are only a few years of data available. Shortening the smoothing parameter means that the portfolio is more responsive to changes in the market, which might seem like an attractive portfolio strategy. However, since the portfolio is likely to be rebalanced whenever the market situation changes, it does come with higher transaction costs. Therefore, in the third case, our aim was to reduce the overall weight changes further by using double smoothing and setting λ to 12 months.

Table 4.6 Performance results for the tangency strategy with and without weight smoothing versus the MICEX Index

	MICEX index	*Portfolio 1 (single smoothing, λ=1 month)*	*Portfolio 2 (single smoothing, λ=3 months)*	*Portfolio 3 (double smoothing, λ=12 months)*
Cumulative Return	0.638	**1.511**	1.295	1.106
Annualized Return	0.122	**0.288**	0.247	0.211
Annualized Std Dev	**0.156**	0.166	0.185	0.176
Annualized Sharpe (Rf=0.00)	0.781	**1.734**	1.335	1.198
Sortino Ratio (MAR = 0.00)	0.388	**1.112**	0.698	0.675
Information Ratio	-	**1.324**	0.891	0.822
R-Squared	-	**0.320**	0.329	0.544
Skewness	0.249	0.168	-0.259	**0.457**
Excess Kurtosis	1.069	**0.468**	2.242	3.911
Maximum Drawdown	**0.151**	0.177	0.267	**0.151**
Historical VaR (0.95)	-0.057	-0.059	-0.065	**-0.052**
Historical ES (0.95)	-0.076	**-0.073**	-0.103	-0.086
Modified VaR (0.95)	-0.059	**-0.051**	-0.068	-0.055
Modified ES (0.95)	-0.079	-0.072	-0.115	**-0.070**
Average Monthly Turnover	-	0.462	0.294	**0.083**
Annual returns:				
2012	0.093	**0.156**	0.128	0.085
2013	0.061	**0.361**	0.307	0.237
2014	-0.019	**0.114**	-0.026	0.010
2015	0.280	0.389	**0.413**	0.387
2016	0.284	**0.430**	0.426	0.311
2017	-0.061	0.060	0.048	**0.076**

As can be seen, the utilization of asset weight smoothing does reduce the average monthly turnover, but it also decreases the risk-adjusted returns. More precisely, the tangency strategy with lambda set to one month yields the highest annualized Sharpe ratio of 1.73 but also the highest turnover rate of 0.46. In contrast, when using double smoothing with lambda set to 12 months, the annualized Sharpe ratio decreases by about 53 percentage points to 1.20, whereas the average monthly turnover decreases by roughly 38 percentage points to 0.08. The second portfolio strikes a balance between these two extremes: it has the second-highest annualized Sharpe ratio of 1.34, as well as the second-highest turnover of 0.29. Interestingly, increasing the decay time from one to three months decreases the annualized return by about four percentage points from 0.29 to 0.25, but this also increases the annualized standard deviation by about two percentage points from 0.17 to 0.19. Indeed, even in the case of double smoothing, the risk measured by the annualized standard deviation was higher than in the case of single smoothing with lambda set to one month (0.18 and 0.17, respectively). The Sortino ratio and information ratio support the conclusions made based on the annualized Sharpe ratio. In other words, the first portfolio clearly outperforms the other two variants, but the difference in performance is narrower between the second and the third portfolio. For instance, although the turnover decreases by about 21 percentage points when switching from the second portfolio to the third, the Sortino ratio decreases only by about two percentage points.

With respect to other statistics, the first portfolio is the least correlated with the MICEX index in terms of the R^2 (0.32), and it also has the lowest (excess) kurtosis of 0.47. The second portfolio is the only one whose returns are negatively skewed (-0.26), whereas the returns of the third portfolio have the highest (positive) skewness of 0.46. Furthermore, the portfolio with double smoothing suffered the least severe maximum drawdown (MDD) of 0.15 and is followed by the first portfolio (0.18). In terms of the downside risk measures reported, the second portfolio is the riskiest: it has the largest MDD of 0.27 and, for instance, the highest modified Expected Shortfall of -0.12. Here, the first and the third portfolio are closer to each other: for example, they both have roughly the same modified ES of -0.07, and the difference between the values of their modified VaR is only about one percentage point (-0.05 and -0.06, respectively).

Based on the annual returns, the first portfolio outperformed its competing variants each year, excluding 2015 and 2017. It was followed by the second portfolio, thus demonstrating how the increased smoothing effect gradually lowers the yields. During the year of turmoil in 2014, the second portfolio was the only one to produce negative returns (-0.03). It was followed by the second portfolio (0.01), whereas the first one was able to react the fastest to the changes in the market environment, leading to an annual return of 0.11. However, the situation changed in the following year of recovery when the second portfolio had the highest return of 0.41, thus slightly outperforming both the first (0.39) and the third portfolio (0.39). In 2017, the third portfolio had about two percentage points higher annual return than the first variant (0.08 and 0.06, respectively). Despite this, the first portfolio was the only one that managed to outperform the MICEX index every year.

Finally, it is worth highlighting that although the average weight changes can be reduced rather efficiently by utilizing weight smoothing, this approach is not a panacea for concentrated portfolios per se. Indeed, since no upper bounds were set on the weights of individual assets, certain assets can at least temporarily get excessive weights. For instance, when double smoothing was used, Surgutneftegas (pref.) received weights exceeding 50 percent for the first 11 months of the backtest period. Similarly, when lambda was set to one (three), Rosseti had weights exceeding 50 percent for the first three (four) months. Hence, in order to achieve more diversified portfolios, one possibility would be to define portfolio concentration as the Herfindahl–Hirschman Index (HHI) of weights and use the concentration aversion parameter to control how much concentration is penalized (see Bennett 2015a, 12; Peterson et al. 2018, 119). We leave these considerations for future research, and the reader interested in reading more about optimal portfolio diversification is referred, e.g., to Bera and Park (2008), Meucci (2009), and Choueifaty, Froidure, and Reynier (2013).

4.2 Minimizing risk and maximizing expected utility with higher-order comoments

When optimized portfolios take higher-order co-moment estimates as the input parameter, non-normality and the risk of 'error maximization' can be accounted for by utilizing the methodology proposed by Boudt et al. (2015). As we have discussed, the building block of this approach is the assumption of a multifactor model generating the asset returns, meaning that parsimony in estimation is achieved by a balance between generality (multiple factors) and the number of parameters to estimate. Here, our first objective is to minimize modified Expected Shortfall (ES) at the 95 percent confidence level, and the second objective is to maximize the Constant Relative Risk Aversion (CRRA) expected utility with the risk aversion parameter gamma (γ) set either to five or ten, as in Boudt et al. (2015, 230). Considering the latter objective, recall from Equation (3.3.7) that an increase in the second and fourth moment will decrease the objective function, whereas an increase in the gamma value places a bigger penalty on the higher moments. The underlying assumption is that investors have preferences for higher odd and lower even moments, which is why the portfolio choice should not be only a simple trade-off between expected return and volatility (see Martellini & Ziemann 2008, 4, 18).

Furthermore, each portfolio objective depends on the estimated comoments, for which three options are evaluated: the sample estimator (*sample*), the single-factor model-based estimate ($k = 1$), and a three-factor approach ($k = 3$). Again, our asset universe consists of 40 stocks, and the backtesting period extends from August 1, 2012, to September 6, 2017. As the optimized portfolio is held through the subsequent month until a new allocation decision takes place, we obtain roughly five years of monthly out-of-sample data. Note also that for portfolios of size $N = 40$, there are 163 elements to estimate under the single-factor model approach and 271 elements under the three-factor model approach. Correspondingly, it requires estimating 135,710 parameters under the unrestricted approach.

Table 4.7 shows the statistics and stylized facts for the risk minimization and expected utility maximization portfolios versus the equal-weight strategy serving as a simple benchmark. The best portfolio for each statistic is highlighted in bold. In addition, Figure 4.7 depicts the time-varying cumulative returns and drawdowns for the most attractive portfolio solutions from each optimization category versus the benchmark. Based on the annualized Sharpe ratios, these include minimizing ES with $k = 3$ (1.91), maximizing CRRA ($\gamma = 5$) with $k = 3$ (1.65), and maximizing CRRA ($\gamma = 10$) with $k = 3$ (1.68). The main observation is that regardless of whether the sample estimator or the statistical factor approach is used, each of the variants outperforms the equally-weighted portfolio in terms of the risk-adjusted returns, and they are less risky in terms of the downside risk measures given, namely the maximum drawdown (MDD) and modified expected shortfall at the 95 percent level. The optimized portfolios also have larger skewness and lower excess kurtosis compared to the benchmark portfolio.

Secondly, when inspecting Figure 4.7, it can be seen that, unlike the benchmark portfolio, the returns of the selected strategies remained positive during the year of crises in 2014. More precisely, the annual return of the equally-weighted portfolio in 2014 was -0.04. In contrast, the annual returns for the multifactor min ES, multifactor max CRRA ($\gamma = 5$), and multifactor max CRRA ($\gamma = 10$) in 2014 were 0.35, 0.31, and 0.19, respectively. Moreover, in February 2014, the equally-weighted portfolio suffered its greatest loss of 0.28, whereas the maximum drawdown of the best-performing strategies was, on average, 0.16, i.e., roughly 42.8 percent (11.8 percentage points) lower.

When the objective is to minimize downside risk, it can be seen that the utilization of the multifactor model increases the annualized Sharpe ratio by 22.4 percentage points from 1.69 to 1.91 compared to the use of sample estimates of the moments. On the other hand, switching from the single-factor model to the three-factor model causes only a 0.7 percentage point increase in the annualized risk-adjusted return. Thus, although the multifactor approach outperforms the sample estimator variant by producing higher annualized return (0.31 and 0.29, respectively) and lower annualized standard deviation (0.16 and 0.17, respectively), it is closely followed by the single-factor approach with a lower annualized return of 0.30 and practically the same annualized standard deviation of 0.16.

Concerning the distribution characteristics of the tail risk-optimal portfolios, we can see that the sample estimator approach produced returns of the highest skewness (0.21) but also the highest excess kurtosis (6.58). The multifactor approach, in turn, generated portfolio returns that show the second-highest skewness of 0.15 and the lowest excess kurtosis of 5.58. Again, this is a substantial improvement compared to the returns of the equally-weighted portfolio, displaying negative skewness of -1.12 and excess kurtosis of 15.13. Finally, the competing variants are fairly similar in terms of the downside risk measures. For instance, the utilization of the single-factor model reduced the maximum loss only by under two percentage points compared to the sample estimator and multifactor approaches, which both suffered the maximum drawdown of 0.15. The difference in modified ES is also practically negligible. Indeed, each of these three variants has roughly the same expected shortfall of -0.02, which, on the other hand, means that there is a hefty 53.5 percent (2.3 percentage point) reduction in risk compared to the benchmark portfolio.

Let us now continue to the expected utility maximization strategy. Here, the first observation is that regardless of whether the statistical factor model has been used or not, an increase in the value of the risk aversion parameter also leads to an increase in the annualized Sharpe ratios of the optimized portfolios. Secondly, in both risk aversion categories, the three-factor approach yields higher risk-adjusted returns than the sample estimator and single-factor methods, which is in line with the findings made in the previous risk minimization case. It is also worth noting that although the risk profile in terms of volatility is not significantly affected by the value given to the risk aversion coefficient, γ does affect the annualized return depending on the estimation method used for the higher-order comoments. More specifically, when γ increases from five to ten, the return of the sample optimized portfolio increases by 2.7 percentage points, whereas its standard deviation increases only by 0.1 percentage points. The difference in annualized returns stays more robust when one utilizes the statistical factor-based estimates. For example, in the multifactor case, an increase in risk aversion leads to a 0.6 percentage point reduction in return, as well as to a 0.6 percentage point reduction in standard deviation. Therefore, in terms of risk-adjusted returns, the performances of the competing variants start to resemble each other as γ increases.

With respect to the distribution characteristics, we can see that the skewness of the returns increases, and excess kurtosis decreases as the value of the risk aversion parameter increases. However, despite this, the returns of each CRRA portfolio are negatively skewed, and they also show higher kurtosis compared to the return distributions of the risk-optimal portfolios. In terms of downside risk measures, the average maximum drawdown is 0.16 when $\gamma = 5$ and 0.17 when $\gamma = 10$. Hence, the max EU strategy can be considered to be slightly riskier than the min ES strategy, whose average MDD is 0.15. The same also applies in terms of expected shortfall. Indeed, the average modified ES of the CRRA portfolios rises roughly from -0.04 to -0.03 as the risk aversion parameter increases from five to ten. Finally, when comparing the performance between the two portfolios with the highest cumulative returns, namely the multifactor minimum ES (2.89) and multifactor maximum CRRA with $\gamma = 5$ (2.21), it can be concluded that the risk-minimizing strategy outperforms the expected utility-maximizing variant by producing higher risk-adjusted returns while simultaneously leading to a reduced expected shortfall.

We remark that although we have imposed the risk budget and weight limit constraints to increase diversification, we have omitted the considerations of transaction costs in this application for the sake of simplicity. Indeed, the main point has been to demonstrate that by accounting for higher-order moments, one can achieve optimal allocation schemes that reduce a portfolio's downside risk by a clear margin compared to an equally-weighted benchmark. Moreover, when the objective is to minimize the tail risk, estimating higher-order comoments under a multifactor specification instead of sample estimators can substantially improve the portfolio's risk-adjusted performance. Naturally, such a finding should be of practical use in the volatile Russian equity market.

Overall, the effects observed here are in line with the research of Ledoit and Wolf (2003), Martellini and Ziemann (2008; 2009), Jondeau and Rockinger (2006ab; 2012), as well as with Boudt et al. (2015). For instance, Martellini and Ziemann (2008) find that improved estimators outperform the sample estimators on average and in probability, as well as in magnitude. However, although it is possible to increase the efficiency of the estimates and the resulting portfolio weights, one should realize that the substantial reduction in dimensionality for structured estimators comes at the cost of specification error (see Martellini & Ziemann 2008 for further discussion). Therefore, more research is needed to determine the optimal number of factors. As an example, if Independent Component Analysis (ICA) was used, one could follow the procedure proposed by Velicer (1976), where the optimal number of components is determined from the matrix of partial correlations (see also Lassance, DeMiguel & Vrins 2018).

Table 4.7 Performance results of the ES minimization and CRRA maximization strategies versus the equal-weighted strategy

	Equal weights	*Min ES sample*	*Min ES k=1*	*Min ES k=3*	*CRRA γ=5 sample*	*CRRA γ=5 k=1*	*CRRA γ=5 k=3*	*CRRA γ=10 sample*	*CRRA γ=10 k=1*	*CRRA γ=10 k=3*
Cumulative Return	1.534	2.607	2.820	**2.890**	1.739	2.055	**2.212**	2.065	2.068	**2.147**
Annualized Return	0.201	0.288	0.302	**0.307**	0.220	0.246	**0.259**	0.247	0.247	**0.253**
Annualized Std Dev	0.165	0.170	**0.159**	0.161	**0.151**	0.157	0.157	0.152	0.153	**0.151**
Annualized Sharpe (Rf=0.00)	1.220	1.689	1.906	**1.913**	1.450	1.573	**1.651**	1.627	1.621	**1.681**
Skewness	-1.118	**0.207**	0.030	0.152	-0.896	-0.944	**-0.696**	-0.628	**-0.531**	-0.701
Excess Kurtosis	15.130	6.581	5.740	**5.581**	10.880	12.000	**9.082**	**6.585**	8.083	8.618
Maximum Drawdown	0.276	0.150	**0.135**	0.154	**0.147**	0.184	0.155	0.174	**0.161**	0.164
Modified ES (0.95)	-0.043	**-0.019**	-0.022	-0.020	-0.037	-0.039	**-0.035**	-0.032	**-0.031**	-0.033

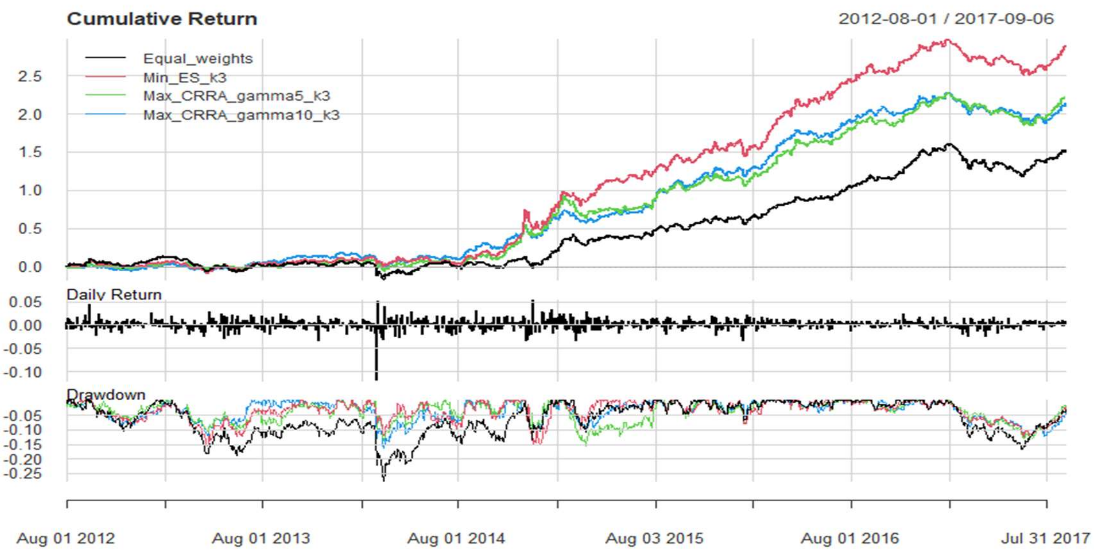

Figure 4.7 Performance summary of the ES minimization and CRRA maximization strategies versus the equal-weighted strategy

With respect to future research, it is also worth mentioning that the framework studied here can be improved further, as is done, for example, in the paper by Boudt, Cornilly, and Verdonck (2020). They propose a minimum distance estimation approach for the higher-order comoments of a multivariate distribution exhibiting a lower-dimensional latent factor structure. A disadvantage of such an approach based on the Nearest Comoment (NC) estimates is that it does not include dynamic behavior in the factor model, which would be needed to accommodate time-variation in the conditional comoments. Alternatively, one could follow Jondeau, Jurczenko, and Rockinger (2018), who describe a statistical technique called Moment Component Analysis (MCA). This method extends Principal Component Analysis (PCA) to higher co-moment tensors and makes it possible to identify the factors driving co-skewness and co-kurtosis structures across a large set of series. Jondeau et al. (2018, 581) find that co-skewness, in particular, conveys useful information about the direction of asset returns for systemic risk measurement and portfolio allocation. Hence, the co-skewness Absorption Ratio (see Kritzman et al. 2011) can be used as a market timing signal.

Finally, as an alternative to portfolio optimization, higher moments could also be utilized in Capital Asset Pricing Models. For instance, Teplova and Shutova (2011) introduce a higher-moment downside framework for conditional and unconditional CAPM in the Russian stock market. They show that incorporating additional risk measures of the third and fourth moments and adopting one-sided risk measures increases the explanatory power for cross-sectional return variations. Similarly, Ranaldo and Favre (2005) extend the two-moment CAPM to a higher-moment model to accommodate coskewness and cokurtosis in the context of pricing hedge fund investments.

4.3 Optimal portfolios based on the relative ranking of expected returns

This subchapter studies the performance of portfolios formed from asset sorts. To express views on the relative order of the returns, we use both the Almgren-Chriss framework (termed AC ranking) and the Fully Flexible Views framework (termed Meucci ranking). Recall that the assets are ranked according to a momentum view based on the previous 21 days. Hence, we take as given that the momentum hypothesis holds and assume that the assets with the highest return will continue to outperform. Our asset universe consists of the nine ruble-denominated sectoral indices, and the out-of-sample evaluation period extends from August 1, 2008, to September 7, 2017.

Table 4.8 shows the performance and downside risk metrics and statistical test results for the relative asset return ranking strategies, as well as for the benchmarks, namely the

MICEX price index, the equal-weight strategy, and the classic quadratic utility (QU) maximization strategy. The most attractive portfolio for each statistic is highlighted in bold. To study the strategies' robustness, we also run the optimizations with the proportional transaction cost constraint, whose value was set to half percent (50 basis points). Recall that this constraint penalizes for changing the asset weights, and the optimal portfolio solutions gradually approach the equal-weighted portfolio as the costs increase. Figure 4.8 illustrates the dynamic behavior of the equal-weight, maximum QU, Meucci ranking, and AC ranking strategies in terms of cumulative returns and drawdowns.

Note that in Table 4.8, we have calculated the single-factor (CAPM) alpha, beta, and R-squared using the MICEX index as a proxy for the market. The risk-free rate is assumed to be zero percent. The alpha intercept measures the amount that the investment strategy has returned in comparison to the benchmark, which is why a high alpha is preferred (see also Jensen 1968, 393–394). Beta, in turn, as a measure of the relative volatility or systematic risk, provides us information on whether a portfolio moves in the same direction as the rest of the market. R-squared is the proportion of a portfolio's movements that can be explained by movements in the MICEX index. Here, a high R-squared can be seen as a bad sign, indicating that the allocation strategy is not adding sufficient value relative to the benchmark. Finally, we utilize the one-sample Wilcoxon signed-rank test to determine whether the median of the returns is equal to zero. This is a non-parametric alternative to a one-sample t-test when the data cannot be assumed to be normally distributed. Thus, the null hypothesis is $H_0: m = 0$, and the corresponding alternative hypothesis is $H_a: m > 0$ (i.e., true location is greater than zero). Statistically significant values at a five percent level are denoted by an asterisk, and the p-values are shown in parentheses.

Several conclusions can be made from these results. Firstly, each optimization variant outperforms the MICEX index and the equally-weighted portfolio in terms of annualized returns and Sharpe ratios, even when the transaction costs (TCs) are included. Secondly, both the benchmarks and the ranking portfolios produced returns with high excess kurtosis and high losses measured by maximum drawdown (MDD). Indeed, caused by the global financial crisis at the turn of 2008 and 2009, the average MDD was about 0.70 among each variant. More precisely, the MICEX index had the lowest MDD of 0.657, whereas the AC ranking portfolio had the highest MDD of 0.714 and 0.736 with and without the TC constraint, respectively. Thirdly, the other downside risk measures reported lead to conflicting conclusions about the riskiness of the competing variants. In other words, although the MICEX index and the portfolios are roughly equally risky in terms of historical VaR and ES at a 95 percent level, the inclusion of transaction costs reduces the downside risk of the optimized portfolios in terms of modified VaR and ES, except in the case of AC ranking. Finally, although the positive alphas produced by the optimized portfolios are not statistically significant at a five percent level, their median returns are significantly greater than zero.

From Figure 4.8, it can be seen that the Meucci ranking strategy closely follows the classical maximum quadratic utility strategy and outperforms it in terms of cumulative returns by 22 percentage points (1.37 and 1.15, respectively). The AC ranking strategy, in turn, follows more closely the equal-weight strategy but underperforms it until the beginning of 2015. After that, the performance of the AC approach improves considerably, and its risk in terms of drawdown is reduced from -0.4 to below -0.2. By the end of August 2017, the AC ranking strategy outperforms each competing variant as its cumulative return reaches the final value of 2.08.

Having established that in terms of risk-adjusted returns, the Almgren-Chriss method outperforms the Fully Flexible Views framework, whereas the latter outperforms the sample optimized maximum quadratic utility strategy, let us now analyze more closely how the inclusion of proportional transaction costs affects the performance of these competing approaches. Although the overall order remains the same, it can be noticed that the annualized returns of the AC ranking strategy decrease by 2.5 percentage points from 0.13 to 0.11, whereas there is an economically more significant drop of 3.9 percentage points from 0.10 to 0.06 in the case of Meucci ranking, as well as a 3.5 percentage point drop from 0.09 to 0.05 in the case of sample optimized portfolio. Correspondingly, the annualized Sharpe ratio falls by 6.6 percentage points from 0.47 to 0.40 in the case of AC ranking, 19.4 percentage points from 0.46 to 0.27 in the case of Meucci ranking, and 17.7 percentage points from 0.41 to 0.23 in the sample optimization case.

Interestingly, the transaction costs decrease the annualized standard deviation of the AC ranking strategy by 1.5 percentage points, which is in contrast to the 1.1 and 1.6 percentage point increase in the Meucci ranking and maximum QU approaches, respectively. Despite this, the returns of the AC ranking strategy remain the most volatile. This is also confirmed by the fact that the AC approach has the highest (and statistically significant) betas compared to the other two competing variants. Furthermore, when the transaction costs are not included, the AC ranking also has the highest R-squared of 0.54. However, when the costs are added, the Meucci ranking becomes the most correlated with the MICEX index, as it now has the highest R-squared of 0.66. In other words, the transaction costs increase the optimized portfolios' betas and make them more correlated with the MICEX index, excluding the AC ranking strategy, which can be considered to be a more robust approach.

With respect to the distribution characteristics, the AC ranking portfolio is the only one with positively skewed returns (0.18) when the transaction costs are not included. However, with the inclusion of costs, its skewness decreases to zero, which is in contrast to the other two optimized portfolios. For example, the skewness of the returns produced by the Meucci ranking strategy increases from -0.86 to 0.06 as the proportional transaction costs are taken into account. On the other hand, the TC constraint also reduces the excess kurtosis in the case of AC ranking, whereas the opposite is true for the maximum QU and Meucci ranking strategies.

Considering the estimated downside risk measures, we can see that although the differences between the optimized portfolios are not immense, they lead to somewhat conflicting conclusions, and the imposition of the TC constraint further complicates the situation. For example, adding the constraint causes a slight increase in the maximum drawdown of the sample optimized and Meucci ranking portfolios, unlike in the AC ranking case. Furthermore, when the trading costs are not included, both the sample optimized portfolio and Meucci ranking portfolio are the least risky in terms of historical ES (-0.03), whereas the AC ranking portfolio has the lowest risk in terms of modified ES (-0.02). However, as the proportional costs are included, the tail risk measured by the modified ES is reduced by 3.7 (2.9) percentage points in the case of the Meucci ranking (sample optimized) portfolio. Again, this is in contrast to the AC ranking strategy, whose risk remains practically unchanged with and without the presence of transaction costs.

Table 4.8 Performance results, risk metrics, and statistical tests for the relative asset return ranking strategies versus benchmarks

	MICEX index	*Equal weights*	*Max QU sample*	*Meucci ranking*	*AC ranking*	*Max QU sample (TC=50 bp)*	*Meucci ranking (TC=50 bp)*	*AC ranking (TC=50 bp)*
Cumulative Return	0.344	**0.577**	1.152	1.373	**2.076**	0.608	0.712	**1.519**
Annualized Return	0.033	**0.052**	0.089	0.101	**0.133**	0.054	0.062	**0.108**
Annualized Std Dev	0.335	**0.252**	**0.216**	0.218	0.283	0.232	**0.229**	0.268
Annualized Sharpe (Rf=0.00)	0.100	**0.206**	0.411	0.463	**0.470**	0.234	0.269	**0.404**
Skewness	**0.857**	-0.242	-0.608	-0.857	**0.182**	**0.129**	0.064	-0.004
Excess Kurtosis	30.154	**29.618**	**13.350**	14.439	24.420	37.345	38.570	**17.902**
Maximum Drawdown	**0.657**	0.685	0.695	**0.680**	0.736	**0.707**	0.710	0.714
Historical VaR (0.95)	-0.027	**-0.023**	-0.020	**-0.019**	-0.025	**-0.021**	**-0.021**	-0.023
Historical ES (0.95)	-0.049	**-0.040**	**-0.034**	**-0.034**	-0.042	-0.037	**-0.036**	-0.040
Modified VaR (0.95)	**-0.016**	-0.017	-0.021	-0.021	**-0.019**	**-0.012**	**-0.012**	-0.021
Modified ES (0.95)	**-0.016**	-0.017	-0.041	-0.049	**-0.019**	**-0.012**	**-0.012**	-0.021
CAPM alpha	-	0.00008	0.00027	0.00031	**0.00044**	0.00012	0.00015	**0.00034**
p-value	-	(0.511)	(0.181)	(0.120)	(0.085)	(0.509)	(0.409)	(0.151)
CAPM beta	-	0.697*	**0.455***	0.466*	0.624*	**0.555***	**0.555***	0.593*
p-value	-	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
R-squared	-	0.858	**0.496**	0.512	0.544	0.644	0.657	**0.549**
V-statistic	1322887	1407625*	1404484*	1407944*	1389680*	1410303*	1412783*	1382534*
p-value	(0.077)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)

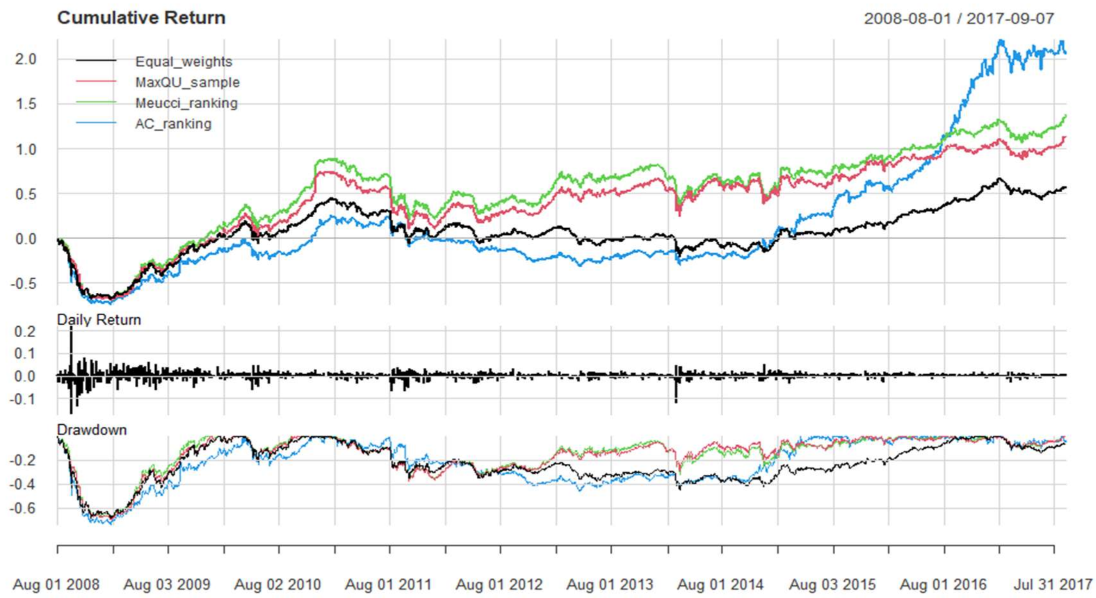

Figure 4.8 Performance summary of the relative asset return ranking strategies versus the quadratic utility maximization and equal-weighted strategies

In conclusion, we have now seen that a strategy based on ordering information and, more specifically, on momentum sorting can achieve economically and statistically significant gains in the Russian stock market. Moreover, the AC framework managed to outperform the Fully Flexible Views framework in terms of risk-adjusted returns, although its downside risk profile may not be quite as attractive in the presence of transaction costs. Therefore, our findings provide a counter-argument to Meucci et al. (2014, 15), who argue that the 'centroid' approach by Almgren and Chriss (2004) presents at least two problems. First, it does not depend on the observed empirical data: two completely different sets of assets with the same relative rankings give rise to the same expected returns. Second, the centroid approach does not alter the volatilities. On the other hand, the Fully Flexible Views framework is more sensitive to parameter uncertainty since it remains as close as possible to the empirical observations summarized by the sample means (Meucci, Ardia & Keel 2013, 11). Thus, the ranking signals (inequalities) have more impact on the model's behavior. This means that one should be able to improve the performance of the Entropy Pooling algorithm with accurate time-varying return (or volatility) forecasts. We return to this issue in the next subchapter. Alternatively, one could follow Meucci et al. (2014, 14) and reorder the stocks using, say, a momentum indicator estimated by exponentially weighted moving averages.

Despite the criticism presented by Meucci et al. (2014), it is also worth emphasizing that the AC algorithm is highly robust to information loss and can provide a significant return on risk, even if the degradation of the information about the order of the expected returns is high (see Almgren & Chriss 2004, 51–55 for the simulation results). Therefore, although a decrease in the quality of a portfolio manager's sort impairs investment performance, perfect knowledge is not required for the AC framework to be economically useful. Rather, a proposed ordering can be expected to be correlated with the true ordering of expected returns. Note, however, that in order to further reduce the estimation error related to the AC ranking method, one could follow, e.g., the approach proposed by Nguyen and Lo (2012). They present a robust ranking model in which the ranking is allowed to be in an uncertainty set. Thus, in this setting, one needs to derive a weight vector that maximizes some generic objective function for the worst realization of the ranking.

With respect to future research, we also remark that although the momentum sorting strategy seems to be able to capture the dynamics of the Russian equity market surprisingly well, a combination of reversal and momentum rules might sometimes work better than each strategy by itself (Chan 2013, 140). For instance, Asness et al. (2013, 945) report that combining these two return strategies increases the Sharpe ratio and outperforms either value or momentum by itself in different markets. Hence, it would be interesting to see whether this 'portfolios from sorts' approach could be modified so that depending on regime switches, it would choose between the value and momentum strategies.

On the other hand, our findings of the profitability of a strategy exploiting short-term momentum are in line with the previous studies. For example, Rouwenhorst (1998b) finds that emerging market stocks exhibit momentum, small stocks outperform large stocks, and value stocks outperform growth stocks. He argues that the return premiums in emerging markets do not simply reflect compensation for liquidity, measured by share turnover. Similarly, Van der Hat et al. (2003, 105, 130) report a significant premium for momentum effect in emerging markets and confirm the profitability of trading strategies based on momentum in these markets. Rockinger and Urga (2000), in turn, test the efficiency of the Russian stock market and conclude that market returns are highly correlated, indicating significant persistency in returns and predictability. Moreover, Hayo and Kutan (2005, 381, 387) find that Russian stock returns are predictable using the previous day's returns, and the degree of persistence is larger than in the case of bond returns. Finally, McGowan and Ibrihim (2009, 29) discover that there is a momentum effect for a three-day 'weekend' for the Russian Trading System (RTS) index and that returns for these three days from Thursday to Monday are persistent.

Kinnunen and Martikainen (2017, 2541) provide further support for the existence of this anomaly by noting that the short-term persistence in equity returns can be related to autocorrelation in returns and cross-autocorrelation among equity returns. Especially in emerging markets, serial correlations and lead-lag relations among stocks are both highlighted. Balsara et al. (2006), in turn, propose a plausible explanation for the momentum profits with a diffusion model. They argue that whereas high trading volume and price volatility are observable factors responsible for a high speed of information dissemination, the low perceived reliability and usefulness of the news is an even more important, albeit unobservable, factor that is responsible for the low receptivity of investors to assimilate this information. Both factors considered jointly result in a low overall rate of information diffusion for high volume and high volatility stocks, leading to higher momentum profits. (Balsara et al. 2006, 407, 421; see also Bernard & Thomas 1989; Hong & Stein 1999.)

Therefore, if we assume that it will be possible to keep exploiting the predictability of the Russian equity market also in the future, one might wish to consider ways to enhance the expected returns of momentum strategies. For instance, in a recent paper, Gulen and Petkova (2018) study the absolute strength strategy, which identifies winners and losers by comparing the stocks' recent performance with the returns of all stocks observed in all periods so far. Yang and Zhang (2019), in turn, find that removing stocks with extreme absolute strength from typical momentum portfolios can improve their performance. However, as has been observed in the US markets, it is also possible that momentum profits will eventually become insignificant. This may be caused, e.g., by a change in the market state (see Lehmann 1990; Cooper, Gutierrez & Hameed 2004; Tokat & Wicas 2007) or improved investor learning and the spread of active trading strategies implemented by arbitrageurs of momentum profits (see Bhattacharya, Kumar & Sonaer 2012).

4.4 Tactical asset allocation with the BL, COP, and EP models

In this subchapter, we study the empirical results obtained by utilizing the Black–Litterman (BL), Copula Opinion Pooling (COP), and Entropy Pooling (EP) portfolio optimization approaches. We start with the BL and COP frameworks and then proceed to examine the EP strategy. For our first analysis, we use the total return indices of four constituents of the Transport sector, namely Aeroflot (AFLT), Fesco (FESH), Novorossiysk Commercial Sea Port (NMTP), and TransContainer (TRCN). The sample starts in July 2011 and ends in September 2017, comprising a total of 75 month-end observations. The first 53 entries are used as a subsample for unit root and co-integration testing, and the remaining data is preserved for executing the backtest.

The degree of integration is determined by applying the Augmented Dickey-Fuller (ADF) and Elliott–Rothenberg–Stock (ERS) tests to the level and first differences of the series. The outcome of these tests is shown in Table 4.9. The critical values for the tests at the five percent level are given as -2.89 (with constant) and -1.95 (without constant) (see Pfaff 2016, 307; Dickey & Fuller 1979; 1981; Elliott, Rothenberg & Stock 1996). Therefore, it can be concluded that, based on the ADF test, all series are difference stationary. Note that in the case of the ERS test, the maximum lag length is set to four by default (Pfaff et al. 2016, 46), and the power of this test can be improved if insignificant intermediate lags are discarded (see Enders & Liu 2014).

Table 4.9 Test statistics for unit root tests

Unit roots	*ADF*		*ERS*	
	Level	Diff	Level	Diff
AFLT	-2.51	-5.26	-1.22	-0.80
FESH	-1.15	-6.48	0.23	-0.71
NMTP	-1.43	-4.66	-1.66	-1.18
TRCN	-1.22	-4.95	-1.42	-1.67

Next, the subsample is fitted to a Vector Error Correction Model (VECM). We use the default lag order of two (expressed in level form) and the transitory specification of the VECM (see Pfaff et al. 2016, 10–12; Pfaff & Stigler 2018, 49–50 for more technical details). The results of the maximum eigenvalue test are shown in Table 4.10. The null hypothesis of no long-run relationship is rejected, and hence the four constituents of the transport sector are co-integrated with rank one. This model specification will be maintained throughout the back-test.

Table 4.10 Results of maximum eigenvalue test for VECM

Co-integration	*Max. eigenvalue*	*Critical values*		
	Statistic	**10%**	**5%**	**1%**
$r \leq 3$	1.97	6.50	8.18	11.65
$r \leq 2$	5.22	12.91	14.90	19.19
$r \leq 1$	**7.43**	18.90	21.07	25.75
$r = 0$	30.69	24.78	27.14	32.14

Having established the structure of the working model, we now generate views by computing the one-step-ahead forecasts and deducting the return expectations from them. As the forecasting rule, we require that only those forecasts considered sufficiently large in absolute terms will enter into a view for a particular company. Thus, we extract the subsample for recursive estimation of the VECM and use the last row of this observation to determine whether a directional view warrants a certain return for a given confidence level. The confidence interval is set to 0.5, and the vector of return expectation is initialized to zero for all four companies. Then, a positive (negative) view on a company is given when the lower (upper) bound for the forecast is above (below) the latest observed value of the total return index. Therefore, the absolute return should be at least as high as the return expectation with $1 - 0.5/2$ confidence. (c.f. Pfaff 2016, 309–311.)

With respect to estimating the parameters of the prior distribution, it is worth noting that the confidences in the views can take any value because the sample variances of the returns will replace them in the estimation of the posterior distribution (i.e., Ω/τ is set to $\kappa P^T \Sigma P$ with $\tau = \kappa = 1$; see Gochez 2015; Gochez et al. 2020, 4). The optimization objective is then to determine the solutions of the maximum Sharpe ratio portfolios derived from the estimates for the parameters of the prior and posterior distributions, respectively. Finally, the weights are allowed to vary in the interval $[-0.8, 0.8]$, the seed wealth is set to 100 monetary units, and the wealth trajectory of a naïve (i.e., equal-weighted) allocation is used as a benchmark.

The paths of the portfolio equities and the relative wealth shares between the Black-Litterman -based portfolio allocations vis-á-vis those determined from the prior distribution and the equal-weighted allocations are displayed in Figure 4.9. In September 2017, the end values of the BL, prior, and equal-weighted portfolios were 317.82, 286.35, and 267.16, respectively. Thus, the value of the BL portfolio more than doubled from November 2015 to September 2017, and it outperformed the 'prior' portfolio by 10.99 percent and the equally-weighted portfolio by 18.96 percent. Similarly, based on the relative performance chart, it can be seen that the BL strategy consistently yielded the highest equity values. Indeed, after the first month of the backtesting period, the BL approach outperformed the equal-weight strategy every month, and, for example, in September 2016, the outperformance was nearly 70 percent. Compared to the allocation based on the prior distribution, the outperformance was more stable and also more moderate. Nevertheless, after the first three months, the BL approach managed to add value even in relation to this variant, and the highest outperformance of about 20 percent was achieved in April 2016.

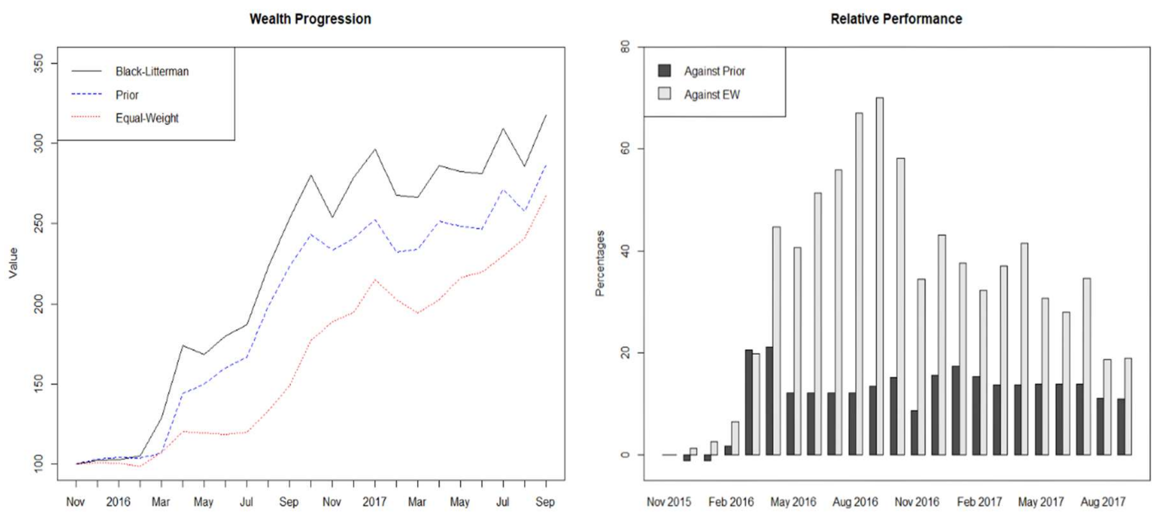

Figure 4.9 Wealth trajectories and relative performance: Equity for BL, prior, and equal-weighted portfolios

Furthermore, the distributions of the allocations obtained from the prior and posterior parameter estimates are depicted as boxplots in Figure 4.10. As can be seen, the weights applied to the total return indices according to the prior distribution are more concentrated compared to the portfolio solution derived from the Black-Litterman posterior distribution. Qualitatively, only short positions would have been taken in the former approach with respect to Fesco's stock, and only long positions for the remaining three stocks. On the contrary, the allocations according to the Black-Litterman posterior span a much wider interval. For example, a long position for Fesco's stock and short positions for Aeroflot's stock would have been entered during the backtest. This reflects the impact of the diverging return expectations between the prior and posterior estimates (Pfaff 2016, 313).

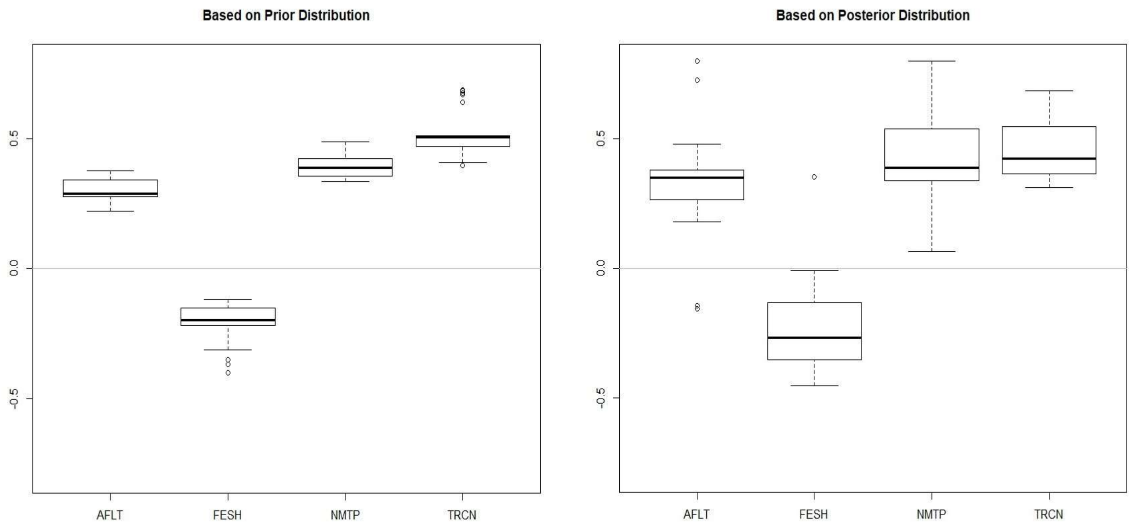

Figure 4.10 Boxplot of weights based on the prior and posterior distributions

We will now continue the previous example by employing the latest return forecasts in the copula opinion pooling framework. In line with Pfaff (2016, 313–314), we assume a multivariate skewed Student's t distribution for the asset returns. For the backtest period (23 months in total), return views different from zero at the 50 percent confidence level were recovered for the total return indices of Fesco (2.59%) and Novorossiysk Commercial Sea Port (1.37%). There were no views for Aeroflot nor TransContainer (i.e., their return estimates were 0.00%). Hence, the pick matrix consists of two rows and four columns and is filled with ones at the relevant row entries for these total return indices. Furthermore, we use the return variances as a measure of dispersion and assume the point forecasts to be normally distributed.

So far, we have defined the specifications for the prior and the views distributions. Next, we assign the copula views and obtain the posterior distribution by the pooling of the random variates thereof. In order to compute the simulated random values of the posterior distribution, we set the simulation size to 10,000 random draws. These simulated random variates are then utilized to derive location and dispersion estimates. Figure 4.11 shows the density plots for the prior and posterior distributions for each of the four companies considered. The prior and posterior densities for the total return indices of Aeroflot and TransContainer pretty much overlap, given that no views have been expressed for these equities. Considering the other two companies, the difference in the shape of the prior and posterior density is the most visible for the total return index of Fesco. This observation primarily reflects differences between the sample means and the stated return forecasts.

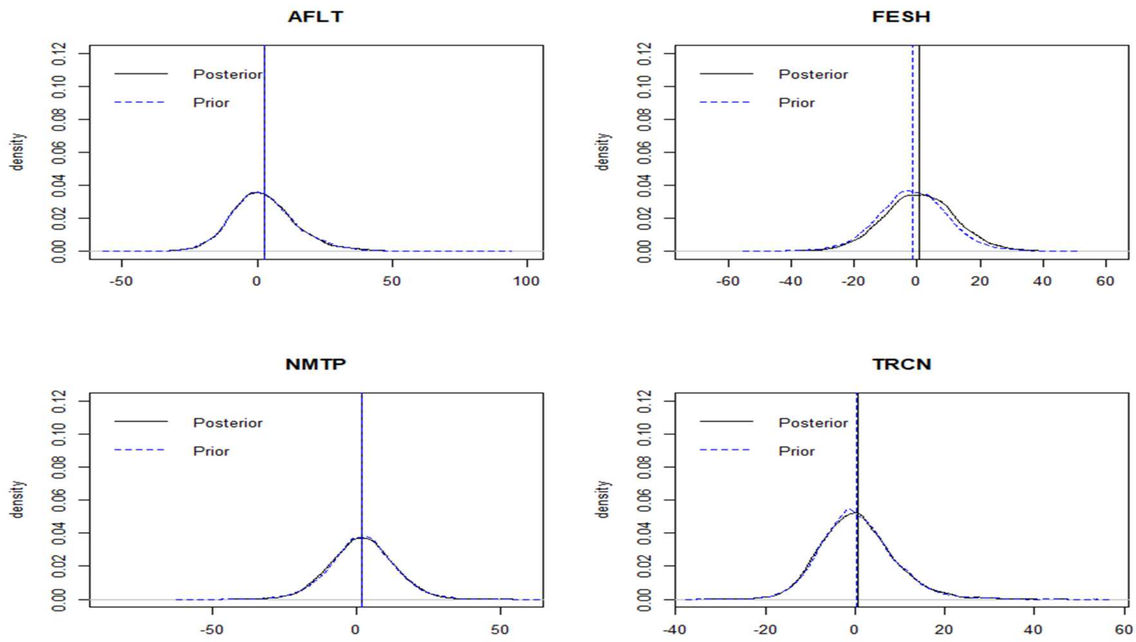

Figure 4.11 Prior and posterior densities for the constituents of the transport sector

Table 4.11 shows how the allocations are determined according to the maximum Sharpe ratio objective for the Gaussian and skewed Student's t prior distributions, as well as according to the Black-Litterman and copula opinion pooling models. The purpose here is to demonstrate that quite differing portfolio allocations can result, given the same set of inputs. From a qualitative point of view, allocations for Novorossiysk Commercial Sea Port (NMTP) and TransContainer (TRCN) coincide, but not for Aeroflot (AFLT) and Fesco (FESH). In other words, only long positions would have been entered for NMTP and TRCN, but both long and short positions are indicated for AFLT and FESH. Moreover, the position sizes differ materially between each of the two prior/posterior pairs and across the BL and COP models. We can also see that the BL portfolio is more concentrated and has a greater risk for losses caused by 'short squeeze' compared to the COP portfolio: the maximum (minimum) investment of the former is 80 (-21) percent in NMTP (AFLT), whereas the maximum (minimum) investment of the latter is 46 (-5) percent in AFLT (FESH). Therefore, given that the BL and COP results were derived from the same set of return forecasts, the latter observation shows the importance of appropriately specifying the marginal distributions and dependence structure. These findings are in line with Pfaff (2016, 316–318).

Table 4.11 Comparison of portfolio allocations expressed as percentages

Model	*AFLT*	*FESH*	*NMTP*	*TRCN*
Gauss	36.71	-21.28	43.24	41.34
Skewed Student's t	20.83	9.80	51.12	18.25
BL	-21.29	26.32	79.80	15.16
BLCop	45.84	-5.02	38.67	20.51

In conclusion, we remark that so far, the applicability of the Copula Opinion Pooling concept has not been widely studied. For example, Stein, Füss, and Drobetz (2009) implement the COP model in the context of fixed-income hedge fund strategies and define views by generating different market scenarios. Simonian (2014), in turn, extends COP to accommodate the aggregation of complex opinions, particularly the assignment of confidence levels to market views. Thus, we hope our application has shed more light on this framework and further shown that COP can indeed be a valuable tool in asset allocation, especially in cases where the investable universe consists of a small number of companies operating in the same industrial sector. Finally, concerning copula functions, it is worth mentioning that they have almost entirely been dedicated to the definition of strategies bound to exploit changes in co-movements, even though copula functions can also be successfully applied to both volatility and correlation trading. Hence, they provide a wide arena for future research. The interested reader is referred, e.g., to Luciano and Schoutens (2006), as well as to Cherubini et al. (2012), who present pricing applications in the context of complex multivariate equity derivatives.

Let us next analyze the performance of the Entropy Pooling (EP) approach combined with volatility forecasting. Here, the total return indices of the 40 constituents of the MICEX Index are used as the dataset. Having computed the discrete percentage returns, we collect all dates pertinent to Wednesdays. In total, there are 311 weekly observations, and the backtest period (76 weeks) starts after an initial period of 235 weeks. Hence, our setup is in line with Pfaff (2016, 320); he also uses a moving data window whose size is roughly 75 percent of the total number of observations.

Table 4.12 shows the key performance measures for the maximum Sharpe ratio portfolios based on the EP posterior, market, and normal distributions (the best values are bolded). Correspondingly, Figure 4.12 depicts the wealth progressions of these three tangency portfolios. As can be seen, the allocations according to the entropy pooling model exceeded those based on either the market or normal distributions in terms of the final wealth values achieved. Moreover, this outperformance was achieved in a relatively short period of time. It is thus apparent that the inclusion of a conditional risk measure (i.e., the one-step-ahead predictions for the conditional volatilities) yields a better risk assessment compared to the result of the market distribution without the blended views. Furthermore, when comparing the wealth trajectories based on the skewed Student's t and multivariate normal distributions, we can notice that they both follow a similar path and that the former is slightly below the latter for most of the out-of-sample period. However, the difference becomes negligible by the end of August 2017. Therefore, the skewed Student's t distribution did not manage to capture the return characteristics noticeably better than the base-case allocation.

The same can be observed by inspecting the statistics listed in Table 4.12. More precisely, the strategy based on the entropy pooling approach produces the highest annualized return of 0.314, outperforming thus the market distribution-based portfolio by 28.6 percentage points. On the other hand, the difference in the annualized returns between the market and normal distribution approaches is only 0.4 percentage points (0.028 and 0.024, respectively). In terms of annualized standard deviation, the base-case allocation has the lowest risk of 0.116, whereas the EP posterior and market distribution approaches are almost identical (0.138 and 0.139, respectively). This means that the EP-based portfolio has an excellent Sharpe ratio of 2.27 compared to the normal and market distribution-based variants, which are close to each other (0.21 and 0.20, respectively). With respect to maximum drawdown, the EP approach suffered the least severe loss of 0.09 and is followed by the portfolio based on the normality assumption (0.121). Allocations produced by these two approaches also have equally favorable modified ES of roughly -0.03. Finally, the allocations based on the market distribution yielded the riskiest progression judged by both downside risk measures. Indeed, this approach has the highest MDD of 0.15, as well as the most negative modified ES of -0.045. Overall, our conclusions resemble those of Pfaff (2016, 323), although he finds that the allocations implied by the market distribution outperform the outcome derived under the normality assumption in terms of risk-adjusted returns when the investable asset universe consists of currencies instead of stocks.

Table 4.12 Key performance measures for the tangency portfolios based on the EP posterior, market, and normal distributions

	Entropy Pooling	*Market*	*Normal*
Annualized Return	**0.314**	0.028	0.024
Annualized Std Dev	0.138	0.139	**0.116**
Annualized Sharpe (Rf=0.00)	**2.270**	0.200	0.210
Maximum Drawdown	**0.090**	0.150	0.121
Modified ES (0.95)	-0.034	-0.045	**-0.030**

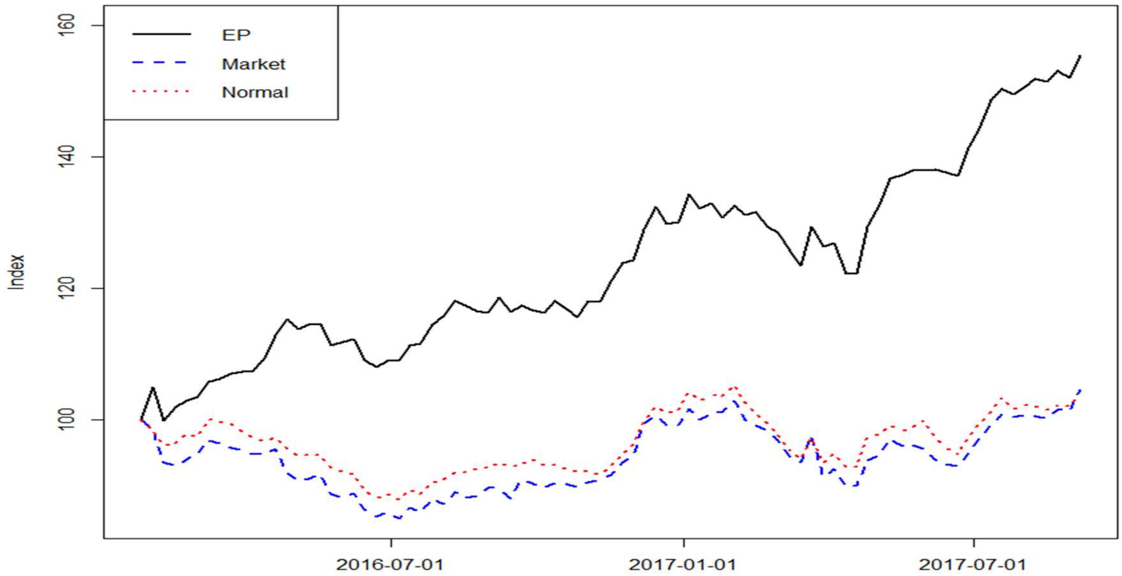

Figure 4.12 Wealth trajectories of tangency portfolios based on the EP posterior, market, and normal distributions

As we have discussed, volatility has a number of regularities that make it suitable for forecasting with time series methods. It certainly seems that the basic GARCH(1,1) model does capture at least some elements of the time evolution of variance. However, there are several reasons why the chosen model may not fit the data well. These include insufficient data (generally, at least 1,000 data points are required), poor initial values for the parameters, as well as persistent seasonality contained in the data. More importantly, the data may not be consistent with the model chosen. In particular, it might be that the fat tails in the return distribution are not due to GARCH effects, so using a normal distribution in the GARCH model does not address the problem completely. (see Sinclair 2013, 56.)

In order to alleviate these concerns, we ran the optimization algorithm several times using different volatility models and conditional distributions included in the *fGarch* (Würtz et al. 2020) and *rugarch* (Ghalanos & Kley 2020) packages. However, as the utilization of these more advanced configurations led to worse risk-adjusted returns than the use of a basic GARCH(1,1) model with normal distribution, we omit reporting these results. Instead, we can better understand why the standard GARCH(1,1) model outperformed competing variants in producing the most accurate one-step-ahead forecasts on average by looking at Table 4.13. Here, we have listed some of the volatility models tested, namely EGARCH(1,1), GJR-GARCH(1,1), and APARCH(1,1), in addition to the benchmark model. To see whether the choice of the conditional density used for the innovations has an effect on the forecasting performance, we also changed the GARCH(1,1) model's normal distribution to Student's *t*, as well as to the skew Generalized Error Distribution (GED).

Considering model selection, Table 4.13 reports the Akaike (AIC), Bayesian (BIC), Shibata (SIC), and Hannan-Quinn (HQIC) information criteria. Furthermore, to make preliminary inferences on the forecast accuracy of the competing variants, we used weekly (log) returns of our 40 assets and computed one-step-ahead forecasts so that 76 data points from the end of the dataset were kept for out-of-sample forecasting, and the number of rolling forecasts to create beyond the first one was set to 75. We then computed the Mean Squared Error (MSE), Mean Absolute Error (MAE), and Directional Accuracy (DAC) of the forecasts versus realized returns for each time series. Finally, we calculated the average values of these fit diagnostics and forecast performance measures and highlighted the most attractive result for each statistic in bold. (see Ghalanos 2017; Ghalanos & Kley 2020 for more technical details.)

As can be seen, the standard GARCH(1,1) model of Bollerslev (1986) with the conditional Student's *t* distribution has, on average, the best fit for the time-series data since it consistently has the smallest values of information criteria. However, the best-fitted model is not necessarily a model that can provide the most accurate volatility forecasts. Indeed, we can find that, although the difference is small, the 'vanilla' GARCH (1,1) model with the conditional normal distribution produces the lowest values for the MSE and MAE loss functions (0.00166 and 0.0284, respectively) and the highest directional accuracy of 0.51. Therefore, we can conclude that this basic configuration is suitable for cases where one needs to produce predictions for dozens of assets using the same model. The more advanced models introduce more parameters to be estimated, which ultimately leads to greater inaccuracies in the forecasts produced. Also, by utilizing the same standard GARCH model as in Pfaff (2016), we have avoided the potential concerns of data snooping and provided further confirmation of the universal profitability of the EP strategy. With respect to future research, it would be interesting to see whether the algorithm could be modified to dynamically choose an optimal volatility model for each time series during every rebalancing period. Unfortunately, however, the issue with such an approach is that it might become infeasible to implement the strategy in practice due to the slow solver, which also often fails to converge when using complex structures.

Table 4.13 Average information criteria and forecast accuracy metrics for alternative volatility modeling configurations

No.	Variance equation	Conditional distribution	AIC	BIC	SIC	HQIC	MSE	MAE	DAC
			Information criterion				*Forecast accuracy (one-step-ahead)*		
1	GARCH(1,1)	Normal	-3.147	-3.088	-3.147	-3.123	**0.001662**	**0.02836**	**0.5102**
2	EGARCH(1,1)	Normal	-3.203	-3.130	-3.204	-3.174	0.001680	0.02849	0.4845
3	GJR-GARCH(1,1)	Normal	-3.170	-3.096	-3.171	-3.141	0.001674	0.02844	0.5043
4	APARCH(1,1)	Normal	-3.187	-3.098	-3.188	-3.151	0.001675	0.02846	0.5003
5	GARCH(1,1)	Student's *t*	**-3.284**	**-3.210**	**-3.285**	**-3.254**	0.001663	0.02837	0.50132
6	GARCH(1,1)	Skew GED	-3.278	-3.190	-3.279	-3.242	0.001674	0.02848	0.5063

Overall, our findings of the average forecast accuracy of different GARCH models are, to some extent, in line with the previous studies. For example, using foreign exchange rate data, Zumbach (2004, 82–83, 85) notes that although there is a vast number of modifications and extensions of the basic univariate GARCH(1,1) process, the quantitative improvements of all these extensions (including long-memory processes) are not spectacular, regardless of the chosen criteria (log-likelihood or forecast error). He also argues that the past volatility generally provides a bad forecast, regardless of the historical depth, with an optimum of around one week. In contrast, Hansen and Lunde (2005) compare the forecast performance of 330 ARCH-type models using both exchange rate and stock return data. They find no evidence that a GARCH(1,1) is outperformed by more sophisticated models in the analysis of exchange rates, whereas the GARCH(1,1) is inferior to models that can accommodate a leverage effect (such as APARCH) in the analysis of returns of a single stock (see also Hamilton & Susmel 1994; Kim & Kon 1994).

Note also that a rather theoretical explanation for why the simple GARCH(1,1) model can produce more attractive performance metrics than more complicated volatility models is presented in Nelson (1990; 1992), as well as in Nelson and Foster (1994; 1995). Indeed, it can be shown that broad classes of ARCH models are surprisingly robust to certain types of misspecification, but this robustness does not extend to all ARCH models, and it breaks down altogether to the point where the process is not well approximated by a diffusion (Nelson 1992, 62). Hence, if the process generating prices is (approximately) a diffusion, then there is so much information about conditional second moments at high frequencies (e.g., daily data) that even a misspecified model can be a consistent filter. In other words, when both observable variables and conditional variances change 'slowly' relative to the sampling interval, then broad classes of ARCH models (even when misspecified) provide continuous-record consistent estimates of the conditional variances. Moreover, the conditional variance estimates generated by the (misspecified) ARCH model convergence in probability to the true conditional variances when the observable variables are recorded at finer and finer intervals. (Nelson 1992, 84; see also Nelson 1994; Andersen et al. 2001, 45.) Nelson and Foster (1995, 304) show further that under suitable conditions, a sequence of misspecified ARCH models may not only be successful at filtering but at forecasting as well.

For this reason, future research might wish to consider using high-frequency intraday data as a potential way to improve the accuracy of the forecasts. Arguments to support such an approach are provided, e.g., in Andersen and Bollerslev (1998), although Martens (2002, 516–517) argues that for the weekly and monthly forecast horizons, it is much less important to utilize an 'intraday' GARCH(1,1) model than for daily horizons. Baillie et al. (1996), in turn, suggest focusing on the proper modeling of long-term volatility dependencies as the forecast horizon lengthens beyond one day. Finally, Hol and Koopman (2002, 21) conclude that in the absence of intraday volatility information, the stochastic volatility model is the preferred model for forecasting volatility. (see also Barndorff-Nielsen & Shephard 2001; Engle 2002; Meddahi 2002; Fleming, Kirby & Ostdiek 2003; Hansen & Lunde 2003; Corsi 2009; Hansen, Huang & Shek 2011 for more discussion on estimation and properties of realized volatility in the context of high-frequency data and intraday returns.)

Still another approach to capture changes in volatility would be to utilize Markov-switching state-space models.[35] For example, a Markov-switching autoregressive conditional heteroscedasticity (SWARCH) model is developed by Hamilton and Susmel (1994), whereas Frijns, Lehnert, and Zwinkels (2011) present a so-called Benchmark Volatility Targeting or BVT-GARCH model. Finally, Jing-rong, Yu-Ke, and Yan (2011) introduce a forecast combination approach where the forecasting results of the GARCH, EGARCH, stochastic volatility, and moving average models are combined based on time-varying weights that can be driven by regime-switching in a latent state variable.

On the other hand, instead of using the GARCH family of models, many alternative methods could be employed to forecast volatility, such as Artificial Neural Networks (ANNs) and genetic algorithms (Sinclair 2013 57). One argument in favor of utilizing algorithms stemming from the artificial intelligence and machine learning community is the fact that the GARCH model and its variants make assumptions on the distribution of the underlying data. Thus, instead of considering alternatives to the Gaussian distribution, one might wish to require no *a priori* assumption. As shown in Kim (2017), a particularly effective algorithm in such a situation is the Support Vector Regression (Drucker et al. 1997). However, GARCH can also be used together with ANNs to create hybrid models

[35] See, e.g., McCulloch and Tsay (1994), Kim, Morley, and Nelson (2001), Tsay (2010), Dai, Zhang, and Zhu (2011), Kritzman, Page, and Turkington (2012), Nguyen, Yin, and Zhang (2014), Ardia et al. (2016), Yang et al. (2016), Ghalanos (2017), Ardia et al. (2018) for more discussions.

(see, e.g., Donaldson & Kamstra 1997; Hajizadeh et al. 2012), which, in turn, can be used as an aid in portfolio optimization (see, e.g., Pojarliev & Polasek 2003).

The final issue with this EP (as well as with the COP) strategy is related to the general framework of mean-variance efficient portfolios. It is well established that the optimized mean-variance portfolio as a complex function of estimated means, volatilities, and correlations of asset returns can 'blow up' when there are tiny errors in any of these estimated input parameters (Ang 2014, 104–104). Thus, instead of choosing the weights to maximize the Sharpe ratio, one might want to consider the risk parity strategy (see Qian 2005), which chooses asset weights proportional to the inverse of variance (or volatility). These low-volatility portfolio selection techniques are discussed, e.g., in Blitz and van Vliet (2007), de Carvalho, Lu and Moulin (2012), and in Frazzini and Pedersen (2014). Similar to our approach, Kim (2017) produces daily volatility forecasts of individual stock returns and finds that especially the general SVR technique can improve the risk-adjusted returns when low-volatility portfolios are formed based on the heuristic approach (see also Chow et al. 2014 and references therein).

Alternatively, future research might consider implementing the entropy pooling approach parametrically under a factor structure (termed Factor Entropy Pooling), as is done in Meucci, Ardia, and Colasante (2014). The factor structure reduces the dimension of the asset correlation structure, linearizes the parameter space, and selects coordinates such that the optimization target becomes unconstrained. Such a parsimonious structure with limited parameters is an example of shrinkage estimation. This approach also retains the flexibility of entropy pooling: views can be expressed as fully general statements, such as relative rankings, spread views, views on volatilities, correlations, tails, et cetera. It can also be used for stress testing different scenarios, as well as for the estimation of the 'implied returns' (see Black & Litterman 1991a). (Meucci et al. 2014, 3.)

4.5 Portfolio protection strategy

In this final subchapter, we analyze the results of the portfolio protection strategy described in Subchapter 3.6. We assume that the tactical asset allocation part of the portfolio can be allocated to the nine sectoral price indices of the Russian stock market and compare the outcome of this approach with the solution of a simple equal-weighted strategy. In line with Pfaff (2016, 328), we extract the dates that fall on Wednesdays, and if some public holidays coincide with this day of the week, the last observed daily observation is carried forward. Thus, there are 490 weekly observations in total, and the moving window is set to a length of 195 weeks (roughly three years and ten months). The trajectory of the ruble-denominated index values is shown in Figure 4.13.

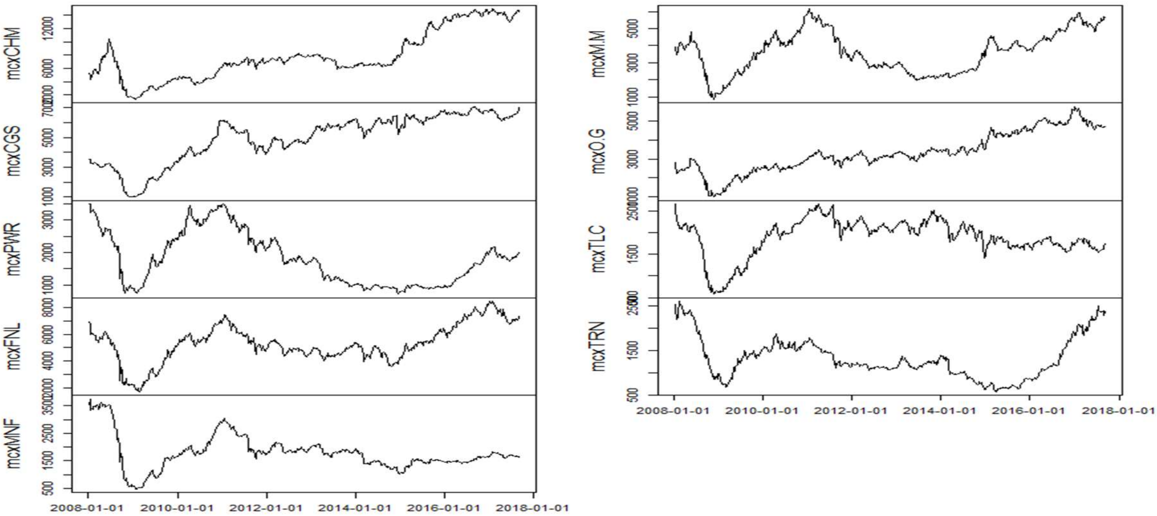

Figure 4.13 Trajectories of sector index values in rubles

Next, the return expectations are taken as the one-step-ahead predictions from the ARIMA variant with the least amount of parameters to be estimated, namely from an ARIMA(1,2,1) model (the series had to be differentiated twice to achieve stationarity). Note that there are at least two reasons for this choice: Firstly, although this simple model may not be optimal for each individual index, the algorithm proposed by Pfaff (2016) allows us to specify only one configuration to be used in forecasting. Secondly, as pointed out in Brockwell and Davis (2002, 169), it is not advantageous from a forecasting point of view to select arbitrarily large values for p and q. Fitting a high-order model will typically result in a small estimated white noise variance. However, as the fitted model is employed for forecasting, the mean squared error of the forecasts will depend not only on the fitted model's white noise variance but also on errors emerging from the estimation of the model's parameters. These will be greater for higher-order models.

The equity curves according to the long-only wealth protection strategy and equal-weighted strategy are shown in Figure 4.14. As we compare the portfolio allocations, we see that the trajectory according to the Tactical Asset Allocation (TAA) strategy leads to almost twice as high a wealth level at the end of the simulation period as the Equal-Weighted (EW) strategy. Moreover, the drawdown witnessed at the beginning of the Ukrainian crisis in 2014 is less pronounced in the case of the former strategy. These two characteristics are also mirrored by the key portfolio performance and risk-related measures, as reported briefly in Table 4.14. Indeed, the TAA portfolio has roughly 11 percentage points higher annualized return than the benchmark (0.20 and 0.09, respectively). In addition, although the former is a more volatile strategy, it has about nine percentage points higher Sharpe ratio than the EW portfolio (0.17 and 0.08, respectively). Finally, both portfolios are relatively similar in terms of downside risk measures, i.e., their Gaussian Value-at-Risk (Expected Shortfall) at the 95 percent level is about three (four) percent.

Table 4.14 Key portfolio performance measures for the protection strategy versus an equally-weighted strategy

	EW long-only	TAA long-only
Annualized return	0.088	0.200
Sharpe ratio	0.080	0.170
VaR (0.95)	0.035	0.032
ES (0.95)	0.044	0.041

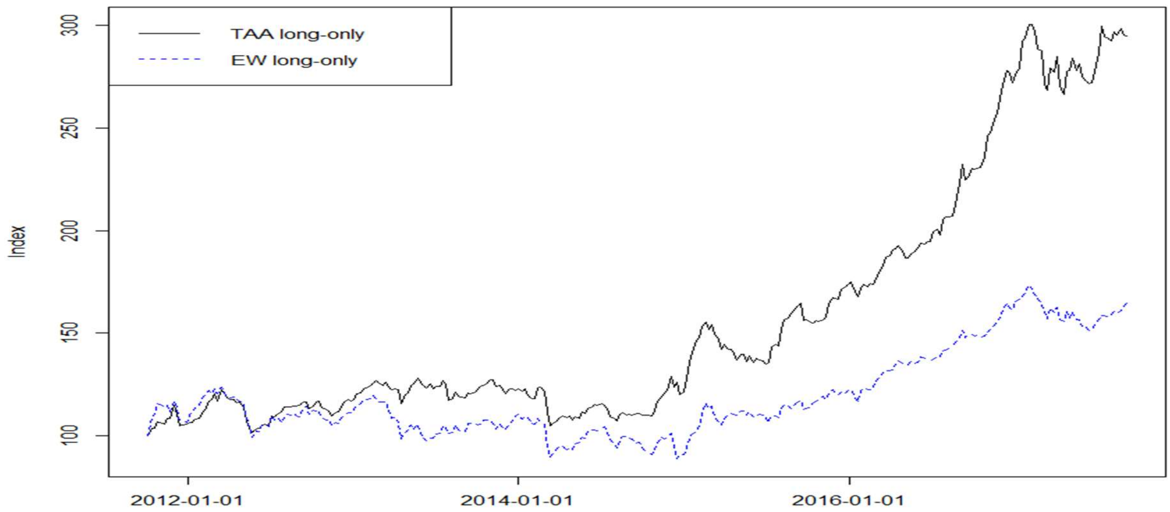

Figure 4.14 Progression of portfolio wealth for the protection strategy versus an equally-weighted strategy

As mentioned earlier, a potential issue with this application is that only one forecasting model has been defined for multiple time series. Also, since returns data tend to be very noisy, long data samples are needed to identify and estimate prediction models (see Timmermann & Granger 2004, 24). Fortunately, however, the choice of the ARIMA(1,2,1) model is supported by Table 4.15. It reports the results of the Diebold-Mariano (DM) test for predictive accuracy of the ARIMA(1,2,1) model versus the best ARIMA model fitted automatically to each univariate time series according to either AIC, AICc, or BIC value (see Hyndman & Khandakar 2008; Hyndman et al. 2018b for further details). The null hypothesis is that the two forecasts have the same accuracy ($H_0 : E(d_t) = 0\ \forall t$), and the alternative hypothesis is that the two forecasts have different levels of accuracy ($H_1 : E(d_t) \neq 0$) (see Diebold & Mariano 1995; Harvey, Leybourne & Newbold 1997). The forecast accuracy of these competing models is compared in two instances: The first test is conducted on in-sample one-step forecasts using thus the entire sample period. The second test is conducted on out-of-sample one-step forecasts so that 295 weekly observations (i.e., $490 - 195$) are withheld from the dataset when fitting, and the one-step forecast errors are then computed for this out-of-sample data (see Diebold 2015 for more discussion on the problems related to this kind of pseudo-out-of-sample analysis). The power used in the loss function is set to two by default in both cases.

Table 4.15 Diebold-Mariano test for predictive accuracy: ARIMA(1,2,1) versus the best model

	Test on in-sample one-step forecasts			*Test on out-of-sample one-step forecasts*		
Sector	Best ARIMA(p,d,q) model	DM test statistic	p-value	Best ARIMA(p,d,q) model	DM test statistic	p-value
CGS	ARIMA(0,1,1)	0.026	0.980	ARIMA(0,1,3)	-2.017	0.045
CHM	ARIMA(1,1,1)	0.516	0.606	ARIMA(0,1,0)	0.944	0.346
FNL	ARIMA(3,1,0)	0.511	0.610	ARIMA(3,1,1)	-0.992	0.322
M&M	ARIMA(3,1,1)	1.486	0.138	ARIMA(0,1,0)	0.345	0.730
MNF	ARIMA(3,1,1)	1.278	0.202	ARIMA(2,2,1)	-1.166	0.244
O&G	ARIMA(2,1,0)	0.041	0.967	ARIMA(3,1,0)	-1.396	0.164
PWR	ARIMA(3,1,0)	0.342	0.732	ARIMA(0,1,0)	0.397	0.692
TLC	ARIMA(2,1,2)	0.606	0.545	ARIMA(4,2,1)	-2.423	0.016
TRN	ARIMA(3,2,3)	0.546	0.585	ARIMA(5,1,0)	-2.139	0.033

The conclusion of this robustness check is two-fold. Firstly, it can be observed that a model that fits the training data well will not necessarily forecast well. Secondly, a perfect fit can always be obtained by using a model with enough parameters. (Hyndman & Athanasopoulos 2014, 51.) Indeed, based on the p-values, we cannot reject the null hypothesis of no difference at the five percent significance level, except in the out-of-sample cases of Consumer Goods & Services (0.05), Telecoms (0.02), and Transportation (0.03). However, additional one-tailed tests revealed that in these three instances, the model fitted by the auto ARIMA method produced significantly less accurate forecasts than the ARIMA(1,2,1) model at the five percent level.

In conclusion, we have now seen that it is possible to construct a dynamic investment strategy that exploits predictability in Russian equity returns. Recall from Subchapter 3.1 that the autocorrelation, albeit quantitatively small, shows that there is information in the past about the future evolution of the returns (Zumbach 2004, 70). More generally, the finding that economic profits can be made by trading on the basis of signals produced from a forecasting model is in line, e.g., with the studies of Lo and MacKinlay (1997), Timmermann and Granger (2004), and Campbell and Thompson (2008). With respect to modeling trading costs, we remark that here, for the sake of simplicity, we have not penalized each week's transaction with a more or less arbitrary amount. Neither have we imposed a proportional transaction constraint, which, as we have seen in earlier demonstrations, would make the TAA portfolio approach the equally-weighted portfolio as the amount of costs increases. Hence, to condense space, we leave further cost analysis for future research.

On a different note, this application has demonstrated that a portfolio protection strategy related to tactical asset allocation can be implemented as an alternative to an option-based hedging strategy. Therefore, it can be considered suitable for risk-averse investors and those market participants who do not wish to include options in their opportunity set due to the higher transaction costs prevalent in the derivatives market. Nevertheless, since wealth protection strategies have been known for a long time,[36] it would be interesting to compare the performance of Pfaff's (2016) linear programming approach with novel option-based strategies. For example, in the market timing strategy proposed by Cherubini et al. (2012, 195–202), a position in protective put options represents portfolio insurance, and the dependence is modeled utilizing the Henriksson–Merton copula (see Henriksson & Merton 1981). Topaloglou, Vladimirou, and Zenios (2011), in turn, develop a stochastic programming model to address in a unified way several interrelated decisions in portfolio management, including optimal portfolio diversification and mitigation of market and currency risks with options.

The final suggestion for future research relates to utilizing more complex methods for the expected return estimation. As noted by Cont (2001, 229), ARMA modeling cannot distinguish between asset returns and white noise, indicating the need for nonlinear dependence measures. Hence, one could follow, for example, Zhang (2003), who proposes a hybrid methodology that combines ARIMA and Artificial Neural Network (ANN) models to take advantage of the unique strength of both models in linear and nonlinear modeling. Based on the author's empirical findings, this kind of combined model can effectively improve forecasting accuracy achieved by either of the models used separately.

Alternatively, one might consider using Hidden Markov Models (HMMs) to predict market regime switches so that wealth could be allocated optimally between, say, bond and stock markets, depending on the current state (e.g., bear or bull). This kind of approach is implemented, for example, by Ang and Bekaert (2002) and Guidolin and Timmermann (2007). Still another approach would be to employ different machine learning-based algorithms, considering that they have been shown to be useful in financial time series forecasting.[37] There are also numerous studies on the applications of neural networks to the portfolio selection problem.[38] From this vast category of methods, especially Long Short-Term Memory (LSTM) networks have recently been a popular tool for forecasting.[39]

---

[36] See, e.g., Leland and Rubinstein (1976), Perold (1986), Perold and Sharpe (1988), Black and Jones (1987), Black and Rouhani (1989), Negrych and Senft (1989), Figlewski, Silber, and Subrahmanyam (1990), Black and Perold (1992), Marshall and Bansal (1992), Aliprantis, Brown, and Werner (2000), Bookstaber and Langsam (2000), Haugh and Lo (2001), Liu and Pan (2003), Prigent (2007), Muck (2010).
[37] See, e.g., Tay and Cao (2001), Pérez-Cruz, Afonso-Rodríguez, and Giner (2003), Gavrishchaka and Banerjee (2006), Ahmed et al. (2010), Yeh, Huang, and Lee (2011).
[38] See, e.g., Lazo, Pacheco, and Vellasco (2002), Fernández and Gómez (2007), Freitas, De Souza, and de Almeida (2009).
[39] See, e.g., Pedregosa et al. (2011), Takeuchi and Lee (2013), Heaton, Polson, and Witte (2016), Krauss, Do, and Huck (2016), Moritz and Zimmermann (2016), Greff et al. (2017), Tsantekidis et al. (2017), Fischer and Krauss (2018), Siami-Namini and Siami-Namin (2018), Xue et al. (2018), Borovykh, Bohte, and Oosterlee (2019), Gu, Kelly, and Xiu (2019), Ghosh, Neufeld, and Sahoo (2020), LeDell et al. (2020).

5 SUMMARY

In this thesis, we have extensively analyzed the performance of quantitative investments strategies in the emerging Russian equity market. The common denominator for these approaches has been that they take into account uncertainties, risks, and dependencies. Moreover, special attention has been paid to the concept known as tactical asset allocation (TAA). In order to alleviate concerns related to backtest overfitting, we have used two datasets with different lengths and sample sizes, namely time series of individual stocks and sectoral indices. Overall, the algorithmic strategies introduced and tested have shown that it is possible to invest profitably and relatively safely in Russia, thus creating new opportunities for international asset managers. We also argue that they have managed to reveal such exploitable predictabilities that may have stayed unknown to most of the readers of this thesis.

For the sake of coherence, our empirical findings and the subsequent conclusions form five separate subchapters. We begin by focusing on classical and robust mean-risk portfolios, as well as on the modeling of transaction costs. The main results of Subchapter 4.1 can be briefly summarized as follows. Firstly, the issue of estimation error related to expected return maximization strategies can be tackled by maximizing the mean return per unit expected shortfall. The performance of this maximum stable tail-adjusted return ratio (STARR) portfolio can be further improved by imposing an equal risk contribution (ERC) constraint requiring that each asset contributes equally to the portfolio downside risk. Secondly, the risk-adjusted returns of minimum variance portfolios can be enhanced by using robust covariance estimators, especially the minimum volume ellipsoid (MVE) estimator. Thirdly, we find that portfolios constructed by utilizing robust optimization techniques outperform the classical Markowitz portfolios in terms of risk-adjusted returns. Also, since these portfolio solutions have low volatility and downside risk, they can be considered most suitable for conservative investors.

Furthermore, we show that the above approaches remain financially viable and outperform the MICEX index even after moderate costs have been taken into account. However, an imposition of a transaction cost constraint means that the optimization algorithm resolves to produce equal-weighted allocations as the proportional costs increase. We end Subchapter 4.1 by focusing on the strategy maximizing the Sharpe ratio (i.e., the tangency strategy) and use it as an example to demonstrate that in addition to a risk budget constraint, optimal portfolios' diversification properties can be altered in various ways to suit the investor's preferences better. These include imposing a market-neutrality (zero-beta) constraint and smoothing the asset weight changes.

In Subchapter 4.2, we follow the methodology proposed by Boudt et al. (2015) and use higher-order moments of asset returns (i.e., covariance, coskewness, and cokurtosis) as input parameters for portfolio optimization. We find that when the objective is to minimize the modified expected shortfall, the inclusion of higher-order comoments under a linear factor model leads to economically significant monetary gains and, more importantly, to a substantial decrease in the downside risk compared to the equally-weighted benchmark allocation. In contrast, the economic benefit and risk reduction are less prominent when the objective is to maximize the fourth-order expansion of the constant relative risk aversion (CRRA) expected utility function.

In Subchapter 4.3, we utilize the framework proposed by Almgren and Chriss (2004; 2005) to mitigate the estimation risk related to the first moments. As our interest shifts towards tactical asset allocation, we construct an algorithm that sorts assets from lowest to highest based on the magnitude of their recent past returns. Thus, assets whose past returns are most positive are deemed most favorable, while assets whose returns are most negative are viewed least favorably. We find that this strategy based on ordering information produces stable allocations, meaning that the inclusion of reasonable transaction costs does not greatly reduce the annualized returns it achieves. Moreover, in terms of risk-adjusted returns, the Almgren-Chriss (AC) approach outperforms the relative ranking approach based on Meucci's (2008b) entropy pooling (EP) framework, as well as the quadratic utility maximization strategy based on sample estimates. Our results also support the underlying assumption behind the AC algorithm stating that, due to the temporary informational inefficiencies, one should be able to exploit short-term momentum effects persistent in the Russian stock market.

In Subchapter 4.4, we first derive one-step-ahead return forecasts from a Vector Error Correction Model (VECM) and find that they can be profitably utilized in the original Black-Litterman (BL) portfolio optimization. However, it arguably does not make much sense to assume that return forecasts for various assets are formed independently of each other, especially when they are derived from a multivariate statistical model. Therefore, we compare the BL model with Meucci's (2006) Copula Opinion Pooling (COP) framework as it relaxes the assumption implying independence between the expressed views. We find that in order to achieve less risky and more diversified portfolio solutions, it is important to specify the marginal distributions and dependence structure appropriately, which is the main benefit of the COP model. As the second case, we re-examine the EP framework using one-step-ahead volatility forecasts as views. We show that this approach outperforms the tangency strategy based on a normal and skewed Student's *t* distribution in terms of risk-adjusted returns. We also conclude that the 'plain vanilla' GARCH(1,1) model is adequate in generating the views for multiple time series.

Finally, in Subchapter 4.5, we analyze the performance of a wealth protection strategy proposed by Pfaff (2016). Here, a linear program is used to form insurance for a portfolio, and the tactical asset allocation between equities and the money market is determined according to the one-step-ahead return forecasts generated by an ARIMA model. Firstly, we find that this dynamic investment strategy can exploit the predictability in the Russian stock market. Secondly, it smoothens drawdowns and produces higher risk-adjusted returns compared to an equal-weight strategy.

All in all, despite the lingering structural issues of the Russian economy, a positive trend can be observed in the main index of the Moscow Exchange. Arguably, it is thus economically reasonable for an international investor to seek ways to benefit from the rising equity returns and even to exploit the inefficiencies of the emerging Russian market. The strategies considered in this thesis have been seen to be profitable, and since they have their own merits with different risk/reward profiles, they can be tailored further to suit the specific needs of both retail and institutional investors. However, it remains to be seen whether the sanctions imposed on Russia will be lifted in the near future or whether the country's isolation from the rest of the world continues. Naturally, these aspects affect the dynamic correlation across Russia's industry sectors, as well as the general feasibility of investment strategies.

Furthermore, it is worth emphasizing that since we have tested these strategies using only equity data, we cannot extrapolate that these techniques will be profitable when the investable universe consists of other asset classes traded on the Moscow Exchange, such as fixed-income or derivatives. On the other hand, we remark that the strategies analyzed have been previously introduced in the context of developed Western financial markets. Therefore, it can be argued that the performance of these approaches is at least to some extent universally consistent, although a careful analysis of the market-specific differences cannot be omitted when planning the practical applicability. Similarly, despite the fact that the reliability of the results is relatively good, the validity may be slightly affected by the optimization algorithm used. For example, the working principles of the differential evolution algorithm is a complex topic that has not been discussed in detail in this thesis. One should also note that we have mainly relied on R because it is a versatile and freely available statistical software. However, its professional use is hampered by the lack of warranty regarding optimization errors.

Another issue we need to be aware of is strategy decay. In other words, strategies that have performed well previously can gradually or even rapidly lose their performance characteristics and thus become unprofitable. Indeed, Russia's financial markets are dynamically adjusting to frequent changes in the political and economic environment, which may reduce the usefulness of past data. Moreover, various (unforeseeable) future shocks, such as crises, wars, and large oil price movements, might completely change the market dynamics. Most strategies are also likely to suffer from diminishing returns sooner or later due to the increasing popularity of high-frequency trading. Since past performance might not be a reliable indicator of future success, it is essential to constantly monitor the market environment and develop new methods to improve the profitability of asset allocation processes.

There are two additional comments to make. Firstly, the problem of portfolio structuring has been viewed from the perspective of an investor operating in Russia and holding assets dominated in rubles. Hence, we have omitted the currency risks faced by an international asset manager. Of course, these foreign exchange-rate risks could be hedged, for example, by utilizing forex futures, options contracts, or other suitable derivatives. Secondly, for the sake of simplicity, we have not considered dividend income taxes for foreigners in Russia. In practice, however, it is important to acknowledge the taxation aspect when calculating the net returns.

Finally, one might ask, what is the best strategy to be employed in the Russian stock market? Unfortunately, there is no clear answer because, ultimately, the choice between different asset allocation strategies depends on an investor's risk preferences. For example, in a bull market environment, she might choose a portfolio that produces the highest risk-adjusted returns in exchange for the risk of slightly larger losses. The risk-optimal approaches might be more appealing in a bear market regime because of their conservatism. Therefore, the strategies we have examined should be primarily seen as tools that investors with different risk appetites can utilize. Also, although this thesis can by no means explore all existing quantitative investment strategies, our aim has been to provide as thorough an exposition as space permits. Still, several areas remain to be investigated, such as strategies utilizing machine learning algorithms, as well as approaches based on maximizing the long-term growth of capital.

REFERENCES

Aas, Kjersti – Czado, Claudia – Frigessi, Arnoldo – Bakken, Henrik (2009) Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics*, Vol. 44 (2), 182–198.

Abad, Pilar – Benito, Sonia – López, Carmen (2014) A comprehensive review of Value at Risk methodologies. *The Spanish Review of Financial Economics*, Vol. 12 (1), 15–32.

Aboura, Sofiane (2015) Leverage vs. feedback: Which effect drives the equity market during stress periods? *Annals of Economics and Statistics*, No. 119/120, 269–288.

Abramov, Alexander – Radygin, Alexander – Chernova, Maria (2017) State-owned enterprises in the Russian market: Ownership structure and their role in the economy. *Russian Journal of Economics*, Vol. 3 (1), 1–23.

Abrosimova, Natalia – Dissanaike, Gishan – Linowski, Dirk (2005) Testing weak-form efficiency of the Russian stock market. EFA 2002 Berlin Meetings Presented Paper. <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=302287>, retrieved 4.1.2018.

Acerbi, Carlo (2002) Spectral measures of risk: A coherent representation of subjective risk aversion. *Journal of Banking & Finance*, Vol. 26 (7), 1505–1518.

Acerbi, Carlo (2007) Coherent measures of risk in everyday market practice. *Quantitative Finance*, Vol. 7 (4), 359–364.

Acerbi, Carlo – Szekely, Balazs (2014) Backtesting expected shortfall: Introducing three model-independent, non-parametric back-test methodologies for expected shortfall. MSCI Inc. <https://www.msci.com/documents/10199/22aa9922-f874-4060-b77a-0f0e267a489b>, retrieved 20.12.2019.

Aelst, Stefan, van – Rousseeuw, Peter (2009) Minimum volume ellipsoid. *Wiley Interdisciplinary Reviews: Computational Statistics*, Vol. 1 (1), 71–82.

Aguilar, Omar – West, Mike (2000) Bayesian dynamic factor models and portfolio allocation. *Journal of Business & Economic Statistics*, Vol. 18 (3), 338–357.

Ahmed, Nesreen K. – Atiya, Amir F. – Gayar, Neamat, El – El-Shishiny, Hisham (2010) An empirical comparison of machine learning models for time series forecasting. *Econometric Reviews*, Vol. 29 (6), 594–621.

Ahn, Dong-Hyun – Boudoukh, Jacob – Richardson, Matthew – Whitelaw, Robert F. (1999) Optimal risk management using options. *The Journal of Finance*, Vol. 54 (1), 359–375.

Aït-Sahalia, Yacine – Brandt, Michael W. (2001) Variable selection for portfolio choice. *The Journal of Finance*, Vol. 56 (4), 1297–1351.

Akindinova, Natalia – Kuzminov, Yaroslav – Yasin, Evgeny (2016) Russia's economy: Before the long transition. *Russian Journal of Economy*, Vol. 2 (1), 219–245.

Alexander, Gordon J. – Baptista, Alexandre M. (2004) A comparison of VaR and CVaR constraints on portfolio selection with the mean-variance model. *Management Science*, Vol 50 (9), 1261–1273.

Aliprantis, C. D. – Brown, D. J. – Werner J. (2000) Minimum-cost portfolio insurance. *Journal of Economic Dynamics and Control*, Vol. 24 (12), 1703–1719.

Allaj, Erindi (2013) The Black-Litterman model: a consistent estimation of the parameter tau. *Financial Markets and Portfolio Management*, Vol. 27 (2), 217–251.

Allen, David E. – Ashraf, Mohammad A. – McAleer, Michael – Powell, Robert J. – Singh, Abhay K. (2013) Financial dependence analysis: Applications of vine copulas. *Statistica Neerlandica*, Vol. 67 (4), 403–435.

Almeida, Carlos – Czado, Claudia – Manner, Hans (2016) Modeling high-dimensional time-varying dependence using dynamic D-vine models. *Applied Stochastic Models in Business and Industry*, Vol. 32 (5), 621–638.

Almgren, Robert – Chriss, Neil (2004) Optimal portfolios from ordering information. 2nd draft. Courant Institute of Mathematical Sciences. New York University. <https://cims.nyu.edu/~almgren/papers/optport.pdf>, retrieved 4.1.2018.

Almgren, Robert – Chriss, Neil (2005) Portfolios from sorts. Draft. Courant Institute of Mathematical Sciences. New York University. < https://cims.nyu.edu/~almgren/papers/sort.pdf>, retrieved 4.1.2018.

Ammann, Manuel – Verhofen, Michael (2006) The effect of market regimes on style allocation. *Financial Markets and Portfolio Management*, Vol. 20 (3), 309–337.

An, Byeong-Je – Ang, Andrew – Collin-Dufresne, Pierre (2015) How often should you take tactical asset allocation decisions? Columbia Business School Research Paper No. 15-27. <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2574512>, retrieved 26.5.2019.

Anatolyev, Stanislav (2008) A 10-year retrospective on the determinants of Russian stock returns. *Research in International Business and Finance*, Vol. 22 (1), 56–67.

Anatolyev, Stanislav – Shakin, Dmitry (2007) Trade intensity in the Russian stock market: Dynamics, distribution and determinants. *Applied Financial Economics*, Vol. 17 (2), 87–104.

Andersen, Torben G. – Bollerslev, Tim (1998) Answering the sceptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, Vol. 39 (4), 885–905.

Andersen, Torben G. – Bollerslev, Tim – Diebold, Francis X. – Ebens, Heiko (2001) The distribution of realized stock return volatility. *Journal of Financial Economics*, Vol. 61 (1), 43–76.

Anderson, Robert M. – Bianchi, Stephen W. – Goldberg, Lisa R. (2012) Will my risk parity strategy outperform? *Financial Analysts Journal*, Vol. 68 (6), 75–93.

Andreev, Vladimir O. – Okunev, Oleg B. – Tinyakov, Sergey Eu. (2010) Extreme value theory and peaks over threshold model: An application to the Russian stock market. *New York Science Journal*, Vol. 3 (6), 102–107.

Andriushin, S. – Burlachkov, V. (2009) Monetary policy and the global financial crisis: Methodological aspects and lessons for Russia. *Problems of Economic Transition*, Vol. 51 (9), 51–67.

Ang, Andrew (2014) *Asset management: A systematic approach to factor investing*. Financial Management Association Survey and Synthesis Series. Oxford University Press, New York, NY.

Ang, Andrew – Bekaert, Geert (2002) International asset allocation with regime switches. *The Review of Financial Studies*, Vol. 15 (4), 1137–1187.

Ankudinov, Andrei B. – Lebedev, Oleg V. (2016) Dividend payouts and company ownership structure amid the global financial crisis: Evidence from Russia. *Post-Communist Economies*, Vol. 28 (3), 384–404.

Ankudinov, Andrei – Ibragimov, Rustam – Lebedev, Oleg (2017) Sanctions and the Russian stock market. *Research in International Business and Finance*, Vol. 40 (1), 150–162.

Annaert, Jan – Osselaer, Sofieke, Van – Verstraete, Bert (2009) Performance evaluation of portfolio insurance strategies using stochastic dominance criteria. *Journal of Banking & Finance*, Vol. 33 (2), 272–280.

Anson, Mark (2004) Strategic versus tactical asset allocation: Beta versus alpha drivers. *The Journal of Portfolio Management*, Vol. 30 (2), 8–22.

Antonacci, Gary (2014) *Dual momentum investing: An innovative strategy for higher returns with lower risk*. McGraw-Hill Education, New York, NY.

Ardia, David – Bluteau, Keven – Boudt, Kris – Catania, Leopoldo (2018) Forecasting risk with Markov-switching GARCH models: A large-scale performance study. *International Journal of Forecasting*, Vol. 34 (4), 733–747.

Ardia, David – Bluteau, Keven – Boudt, Kris – Catania, Leopoldo – Trottier, Denis-Alexandre (2016) Markov–switching GARCH models in R: The MSGARCH package. Social Science Research Network (SSRN). <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2845809>, retrieved 19.7.2019.

Ardia, David – Boudt, Kris – Carl, Peter – Mullen, Katharine M. – Peterson, Brian (2010) Differential evolution (DEoptim) for non-convex portfolio optimization. Working Paper. Munich Personal RePEc Archive (MPRA). <https://mpra.ub.uni-muenchen.de/28187/1/MPRA_paper_28187.pdf>, retrieved 20.8.2020.

Ardia, David – Mullen, Katharine – Peterson, Brian – Ulrich, Joshua – Boudt, Kris (2020) Package 'DEoptim': Global optimization by differential evolution. The Comprehensive R Archive Network (CRAN). <https://cran.r-project.org/web/packages/DEoptim/DEoptim.pdf>, retrieved 1.5.2020.

Arévalo, Andrés – León, Diego – Hernandez, German (2019) Portfolio selection based on hierarchical clustering and inverse-variance weighting. In: *Computational science – ICCS 2019*, eds. João M. F. Rodrigues – Pedro J. S. Cardoso – Jânio Monteiro – Roberto, Lam – Valeria V. Krzhizhanovskaya – Michael H. Lees – Jack J. Dongarra – Peter M.A. Sloot, 315–325. 19th International Conference, Faro, Portugal, June 12–14, 2019, Proceedings, Part III. Lecture Notes in Computer Science, Vol. 11538. Springer Nature Switzerland AG, Cham.

Arnuk, Sal L. – Saluzzi, Joseph C. (2012) *Broken markets: How high frequency trading and predatory practices on Wall Street are destroying investor confidence and your portfolio*. FT Press, Upper Saddle River, NJ.

Arruda, Nelson – Bergeron, Alain – Kritzman, Mark (2021) Optimal currency hedging: Horizon matters. *The Journal of Alternative Investments*, Vol. 23 (4), 122–130.

Artzner, Philippe – Delbaen, Freddy – Eber, Jean-Marc – Heath, David (1999) Coherent measures of risk. *Mathematical Finance*, Vol. 9 (3), 203–228.

Asness, Clifford S. – Frazzini, Andrea – Pedersen, Lasse H. (2012) Leverage aversion and risk parity. *Financial Analysts Journal*, Vol. 68 (1), 47–59.

Asness, Clifford S. – Moskowitz, Tobias J. – Pedersen, Lasse Heje (2013) Value and momentum everywhere. *The Journal of Finance*, Vol. 68 (3), 929–985.

Azzalini, Adelchi – Capitanio, Antonella (2014) *The skew-normal and related families*. IMS Monographs. Cambridge University Press, Cambridge.

Babameto, Elton – Harris, Richard D.F. (2008) Exploiting predictability in the returns to value and momentum investment strategies: A portfolio approach. Working paper. University of Exeter. <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1302772 >, retrieved 28.10.2019.

Babecký, Jan – Komárek, Luboš – Komárková, Zlatuše (2013) Convergence of returns on Chinese and Russian stock markets with world markets: National and sectoral perspectives. *National Institute Economic Review*, Vol. 223 (1), 16–34.

Backtest investment strategies. MATLAB & Simulink Example. The MathWorks, Inc. <https://www.mathworks.com/help/finance/backtest-investment-strategies.html>, retrieved 10.3.2021.

Bagheri, Sajjad – Akbarpour, Hamid Reza (2016) Reinvestigation of the West's sanctions against Russia in the crisis of Ukraine and Russia's reaction. *Procedia Economics and Finance*, Vol. 36 (1), 89–95.

Bailey, David H. – López de Prado, Marcos (2012) Balanced baskets: A new approach to trading and hedging risks. *The Journal of Investment Strategies (Risk Journals)*, Vol. 1 (4), 21–62.

Baillie, Richard T. – Bollerslev, Tim – Mikkelsen, Hans Ole (1996) Fractionally integrated generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, Vol. 74 (1), 3–30.

Balakrishnan, N. – Lai, Chin-Diew (2009) *Continuous bivariate distributions*. 2nd ed. Springer, New York, NY.

Balsara, Nauzer J. – Zheng, Lin – Vidozzi, Andrea – Vidozzi, Luca (2006) Explaining momentum profits with an epidemic diffusion model. *Journal of Economics and Finance*, Vol. 30 (3), 407–422.

Banerjee, Anindya – Dolado, Juan J. – Galbraith, John W. – Hendry David F. (1993) *Co-integration, error correction, and the econometric analysis of non-stationary data*. Advanced Texts in Econometrics. Oxford University Press, New York, NY.

Barberis, Nicholas (2000) Investing for the long run when returns are predictable. *The Journal of Finance*, Vol. 55 (1), 225–264.

Barberis, Nicholas – Shleifer, Andrei – Vishny, Robert (1998) A model of investor sentiment. *Journal of Financial Economics*, Vol. 49 (1), 307–343.

Barberis, Nicholas – Shleifer, Andrei (2003) Style investing. *Journal of Financial Economics*, Vol. 68 (1), 161–199.

Barndorff-Nielsen, Ole E. – Shephard, Neil (2001) Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, Vol. 63 (2), 167–241.

Basel Committee on Banking Supervision (2019) Minimum capital requirements for market risk. Bank for International Settlements (BIS). <https://www.bis.org/bcbs/publ/d457.pdf>, retrieved 1.8.2019.

Bawa, Vijay S. (1978) Safety-first, stochastic dominance, and optimal portfolio choice. *The Journal of Financial and Quantitative Analysis*, Vol. 13 (2), 255–271.

Baz, Jamil – Guo, Helen (2017) An asset allocation primer: Connecting Markowitz, Kelly and risk parity. PIMCO, Quantitative Research. October 2017. <https://global.pimco.com/en-gbl/insights/viewpoints/research/an-asset-allocation-primer-connecting-markowitz-kelly-and-risk-parity>, retrieved 1.8.2021.

Beach, Steven L. – Orlov, Alexei, G. (2007) An application of the Black–Litterman model with EGARCH-M-derived views for international portfolio management. *Financial Markets and Portfolio Management*, Vol. 21 (2), 147–166.

Beder, Tanya Styblo (1995) VAR: Seductive but dangerous. *Financial Analysts Journal*, Vol. 51 (5), 12–24.

Bedford, Tim – Cooke, Roger M. (2001) Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial Intelligence*, Vol. 32 (1), 245–268.

Bedford, Tim – Cooke, Roger M. (2002) Vines: A new graphical model for dependent random variables. *The Annals of Statistics*, Vol. 30 (4), 1031–1068.

Bekaert, Geert – Harvey, Campbell R. (1997) Emerging equity market volatility. *Journal of Financial Economics*, Vol. 43 (1), 29–77.

Bendtsen, Claus (2015) Package 'pso': Particle swarm optimization. The Comprehensive R Archive Network (CRAN). <https://cran.r-project.org/web/packages/pso/pso.pdf>, retrieved 1.5.2020.

Bengtsson, Christoffer – Holst, Jan (2002) On portfolio selection: Improved covariance matrix estimation for Swedish asset returns. Working Paper. CiteSeerX. <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.196.9401&rep=rep1&type=pdf>, retrieved 1.8.2021.

Bennett, Ross (2014) Complex portfolio optimization with PortfolioAnalytics. *Presentation at the R/Finance 2014: Applied Finance with R Conference*, Chicago, IL, USA, May 16, 2014, 1–51. <http://past.rinfinance.com/agenda/2014/workshop/RossBennett.pdf>, retrieved 15.5.2021.

Bennett, Ross (2015a) Introduction to PortfolioAnalytics. The CRAN Package Repository. <https://cran.r-project.org/web/packages/PortfolioAnalytics/vignettes/portfolio_vignette.pdf>, retrieved 12.10.2017.

Bennett, Ross (2015b) Using the ROI solvers with PortfolioAnalytics. The CRAN Package Repository. <https://cran.r-project.org/web/packages/PortfolioAnalytics/vignettes/ROI_vignette.pdf>, retrieved 10.10.2017.

Bennett, Ross (2015c) Advanced moment estimation & optimization with PortfolioAnalytics. R/Finance 2015. <http://rossb34.github.io/PortfolioAnalyticsPresentation2015/#1>, retrieved 19.10.2017.

Bennett, Ross (2015e) Custom moment and objective functions. The CRAN Package Repository. <https://cran.r-project.org/web/packages/PortfolioAnalytics/vignettes/custom_moments_objectives.pdf>, retrieved 19.10.2017.

Ben-Tal, A. – Nemirovski, A. (1998) Robust convex optimization. *Mathematics of Operations Research*, Vol. 23 (4), 769–805.

Bera, Anil K. – Jarque, Carlos M. (1981) An efficient large-sample test for normality of observations and regression residuals. Working Paper in Economics and Econometrics, No. 40. Australian National University, Faculty of Economics and Research School of Social Sciences, Canberra.

Bera, Anil K. – Park, Sung Y. (2008) Optimal portfolio diversification using the maximum entropy principle. *Econometric Reviews*, Vol. 27 (6), 484–512.

Berezinets, Irina – Ilina, Yulia – Smirnov, Marat – Bulatova, Liliia (2017) How does stock market react to dividend surprises? Evidence from emerging markets of India and Russia. *Journal of Asia-Pacific Business*, Vol. 18 (3), 153–179.

Berezinskaya, Olga (2017) Investment drought in the Russian economy: Structural characteristics and turnaround perspectives. *Russian Journal of Economics*, Vol. 3 (1), 71–82.

Bernard, Victor L. – Thomas, Jacob K. (1989) Post-earnings-announcement drift: Delayed price response or risk premium? *Journal of Accounting Research*, Vol. 27 (1), 1–36.

Bertsimas, Dimitris – Brown, David B. – Caramanis, Constantine (2011) Theory and applications of robust optimization. *SIAM Review*, Vol. 53 (3), 464–501.

Best, Michael J. – Grauer, Robert M. (1991) On the sensitivity of mean-variance-efficient portfolios to changes in assets means: Some analytical and computational results. *Review of Financial Studies*, Vol. 4 (1), 315–342.

Best, Michael J. – Grauer, Robert R. (1992) The analytics of sensitivity analysis for mean-variance portfolio problems. *International Review of Financial Analysis*, Vol. 1 (1), 17–37.

Bevan, Andrew – Winkelmann, Kurt (1998) Using the Black-Litterman global asset allocation model: Three years of practical experience. *Fixed Income Research*, Goldman Sachs, 1–19.

Beyer, Hans-Georg – Sendhoff, Bernhard (2007) Robust optimization – A comprehensive survey. *Computer Methods in Applied Mechanics and Engineering*, Vol. 196 (34), 3190–3218.

Bhattacharya, Debarati – Kumar, Raman – Sonaer, Gokhan (2012) Momentum loses its momentum: Implications for market efficiency. Midwest Finance Association 2012 Annual Meetings Paper. <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1928764>, retrieved 1.5.2020.

Biglova, Almira – Jašić, Teo – Rachev, Svetlozar – Fabozzi, Frank J. (2004) Profitability of momentum strategies: Application of novel risk/return ratio stock selection criteria. *Investment Management and Financial Innovations*, Vol. 1 (4), 47–61.

Black, Bernard S. – Love, Inessa – Rachinsky, Andrei (2006) Corporate governance indices and firms' market values: Time series evidence from Russia. *Emerging Markets Review*, Vol. 7 (1), 361–379.

Black, Fischer (1976) Studies of stock price volatility changes. In: *Proceedings of the 1976 Meeting of the Business and Economics Statistics Section*, 177–181. American Statistical Association, Washington, D.C.

Black, Fischer (1993) Estimating expected return. *Financial Analysts Journal*, Vol. 49 (5), 36–38.

Black, Fischer – Jones, Robert W. (1987) Simplifying portfolio insurance. *The Journal of Portfolio Management*, Vol. 14 (1), 48–51.

Black, Fischer – Litterman, Robert (1991a) Asset allocation: Combining investor views with market equilibrium. *The Journal of Fixed Income*, Vol. 1 (2), 7–18.

Black, Fischer – Litterman, Robert (1991b) Global asset allocation with equities, bonds, and currencies. *Fixed Income Research*, Goldman Sachs, 1–44.

Black, Fischer – Litterman, Robert (1992) Global portfolio optimization. *Financial Analysts Journal*, Vol. 48 (5), 28–43.

Black, Fischer – Perold André F. (1992) Theory of constant proportion portfolio insurance. *Journal of Economic Dynamics and Control*, Vol. 16 (3-4), 403–426.

Black, Fischer – Rouhani, Ramine (1989) Constant proportion portfolio insurance and the synthetic put option: A comparison. In: *Institutional investor focus on investment management*, ed. Frank J. Fabozzi, 695–708. Ballinger, Cambridge, Massachusetts.

Blitz, David C. – Vliet, Pim, van (2007) The volatility effect: Lower risk without lower return. *The Journal of Portfolio Management*, Vol. 34 (1), 102–113.

Blitz, David C. – Vliet, Pim, van (2008) Global tactical cross-asset allocation: Applying value and momentum across asset classes. *The Journal of Portfolio Management*, Vol. 35 (1), 23–38.

Bodnar, Taras – Okhrin, Yarema – Vitlinskyy, Valdemar – Zabolotskyy, Taras (2018) Determination and estimation of risk aversion coefficients. *Computational Management Science*, Vol. 15 (2), 297–317.

Bollerslev, Tim (1986) Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, Vol. 31 (1), 307–327.

Bollerslev, Tim – Chou, Ray Y. – Kroner, Kenneth F. (1992) ARCH modeling in finance: A review of the theory and empirical evidence. *Journal of Econometrics*, Vol. 52 (1), 5–59.

Bollerslev, Tim – Engle, Robert F. – Nelson, Daniel B. (1994) ARCH models. In: *Handbook of econometrics, vol. 4*, eds. Robert Engle – Daniel McFadden, 2959–3038. North-Holland, Amsterdam.

Bollerslev, Tim – Litvinova, Julia – Tauchen, George (2006) Leverage and volatility feedback effects in high-frequency data. *Journal of Financial Econometrics*, Vol. 4 (3), 353–384.

Bonami, P. – Lejeune, M. A. (2009) An exact solution approach for portfolio optimization problems under stochastic and integer constraints. *Operations Research*, Vol. 57 (3), 650–670.

Bondt, Werner F. M., De – Thaler, Richard (1985) Does the stock market overreact? *The Journal of Finance*, Vol. 40 (3), 793–805.

Bookstaber, Richard – Langsam, Joseph A. (2000) Portfolio insurance trading rules. *The Journal of Futures Markets*, Vol. 20 (1), 41–57.

Borchers, Hans W. (2021) Package 'pracma': Practical numerical math functions. The Comprehensive R Archive Network (CRAN). <https://cran.r-project.org/web/packages/pracma/pracma.pdf>, retrieved 8.5.2021.

Borovykh, Anastasia – Bohte, Sander – Oosterlee, Cornelis W. (2019) Dilated convolutional neural networks for time series forecasting. *Journal of Computational Finance*, Vol. 22 (4), 73–101.

Boudoukh, Jacob – Richardson, Matthew – Thapar, Ashwin – Wang, Franklin (2019) Optimal currency hedging for international equity portfolios. *Financial Analysts Journal – Forthcoming*, 1–19.

Boudt, Kris – Carl, Peter – Peterson, Brian (2010) Vignette: Portfolio optimization with CVaR budgets in PortfolioAnalytics. The Comprehensive R Archive Network (CRAN). <https://cran.r-project.org/web/packages/PortfolioAnalytics/vignettes/risk_budget_optimization.pdf>, retrieved 1.5.2020.

Boudt, Kris – Carl, Peter – Peterson, Brian G. (2013) Asset allocation with conditional value-at-risk budgets. *Journal of Risk*, Vol. 15 (3), 39–68.

Boudt, Kris – Cornilly, Dries – Verdonck, Tim (2020) Nearest comoment estimation with unobserved factors. *Journal of Econometrics*, Vol. 217 (2), 381–397.

Boudt, Kris – Lu, Wanbo – Peeters, Benedict (2015) Higher order comoments of multifactor models and asset allocation. *Finance Research Letters*, Vol. 13 (1), 225–233.

Boudt, Kris – Peterson, Brian – Croux, Christophe (2008) Estimation and decomposition of downside risk for portfolios with non-normal returns. *Journal of Risk*, Vol. 11 (2), 79–103.

Boyd, Stephen – Mueller, Mark T. – O'Donoghue, Brendan – Wang, Yang (2014) Performance bounds and suboptimal policies for multi-period investment. *Foundations and Trends® in Optimization*, Vol. 1 (1), 1–72.

Boyd, Stephen – Vandenberghe, Lieven (2004) *Convex optimization*. Cambridge University Press, New York, NY.

Braga, Maria Debora – Natale, Francesco Paolo (2007) TEV sensitivity to views in Black-Litterman model. 20th Australasian Finance & Banking Conference 2007 Paper. <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1009635>, retrieved 23.5.2019.

Brandt, Michael W. (2005) Portfolio choice problems. In: *Handbook of financial econometrics, vol. 1: Tools and techniques*, eds. Aït-Sahalia, Yacine – Hansen, Lars Peter, 269–336. North-Holland Publishing Company, Amsterdam.

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

Brechmann, Eike Christian – Czado, Claudia (2013) Risk management with high-dimensional vine copulas: An analysis of the Euro Stoxx 50. *Statistics & Risk Modeling*, Vol. 30 (4), 307–342.

Brechmann, Eike Christian – Czado, Claudia (2015) COPAR–multivariate time series modeling using the copula autoregressive model. *Applied Stochastic Models in Business and Industry*, Vol. 31 (4), 495–514.

Breen, William – Glosten, Lawrence R. – Jagannathan, Ravi (1989) Economic significance of predictable variations in stock index returns. *The Journal of Finance*, Vol. 44 (5), 1177–1189.

Broadie, Mark (1993) Computing efficient frontiers using estimated parameters. *Annals of Operations Research*, Vol. 45 (1), 21–58.

Brocato, Joe – Chandy, P. R. (1994) Does market timing really work in the real world? *The Journal of Portfolio Management*, Vol. 20 (2), 39–44.

Brockwell, Peter J. – Davis, Richard A. (2002) *Introduction to time series and forecasting*. 2nd ed. Springer Texts in Statistics. Springer-Verlag, New York, NY.

Brown, Philip – Walter, Terry S. (2013) The CAPM: Theoretical validity, empirical intractability and practical applications. *ABACUS: A Journal of Accounting, Finance and Business Studies*, Vol. 49 (1), 44–50.

Browne, Sid (1999) Reaching goals by a deadline: Digital options and continuous-time active portfolio management. *Advances in Applied Probability*, Vol. 31 (2), 551–577.

Browne, Sid (2000) Can you do better than Kelly in the short run? In: *Finding the edge: Mathematical analysis of casino games*, eds. Olaf Vancura – Judy A. Cornelius – William R. Eadington, 215–231. Institute for the Study of Gambling and Commercial Gaming, University of Nevada, Reno, NV.

Brunnermeier, Markus – Pedersen, Lasse (2005) Predatory trading. *The Journal of Finance*, Vol. 60 (4), 1825–1863.

Brunnermeier, Markus – Pedersen, Lasse (2009) Market liquidity and funding liquidity. *Review of Financial Studies*, Vol. 22 (6), 2201–2238.

Buckley, Neil (2017) Russian agriculture sector flourishes amid sanctions. Financial Times. <https://www.ft.com/content/422a8252-2443-11e7-8691-d5f7e0cd0a16>, retrieved 27.10.2018.

Buhmann, J. M. – Gronskiy, A. Y. – Mihalák, M. – Pröger, T. – Šrámek, R. – Widmayer, P. (2018) Robust optimization in the presence of uncertainty: A generic approach. *Journal of Computer and System Sciences*, Vol. 94 (1), 135–166.

Buklemishev, O. V. – Maliutina, M. S. (1998) Analysis of the Russian stock market's efficiency. *Matekon*, Vol. 34 (1), 21–38.

Butler, John S. – Schachter, Barry (1986) Unbiased estimation of the Black-Scholes formula. *Journal of Financial Economics*, Vol. 15 (1), 341–357.

Campbell, John Y. (1987) Stock returns and the term structure. *Journal of Financial Economics*, Vol. 18 (2), 373–399.

Campbell, John Y. (1996) Understanding risk and return. *Journal of Political Economy*, Vol. 104 (2), 298–345.

Campbell, John Y. – Hentschel, Ludger (1992) No news is good news: An asymmetric model of changing volatility in stock returns. *Journal of Financial Economics*, Vol. 31 (3), 281–318.

Campbell, John Y. – Lo, Andrew W. – MacKinlay, Craig A. (1997) *The econometrics of financial markets*. Princeton University Press, Princeton, NJ.

Campbell, John Y. – Thompson, Samuel B. (2008) Predicting excess stock returns out of sample: Can anything beat the historical average? *The Review of Financial Studies*, Vol. 21 (4), 1509–1531.

Caporale, Guglielmo Maria – Zakirova, Valentina (2017) Calendar anomalies in the Russian stock market. *Russian Journal of Economics*, Vol. 3 (1), 101–108.

Carl, Peter – Peterson, Brian – Boudt, Kris (2010) Business objectives and complex portfolio optimization. R/Finance Workshop 2010. <http://past.rinfinance.com/agenda/2010/Carl+Peterson+Boudt_Tutorial.pdf>, retrieved 25.9.2017.

Carl, Peter – Peterson, Brian G. (2014) PerformanceAnalytics: Charts and tables overview. The CRAN Package Repository. <https://cran.r-project.org/web/packages/PerformanceAnalytics/vignettes/PA-charts.pdf>, retrieved 25.9.2017.

Carvalho, Raul Leote, de – Lu, Xiao – Moulin, Pierre (2012) Demystifying equity risk-based strategies: A simple alpha plus beta description. *The Journal of Portfolio Management*, Vol. 38 (3), 56–70.

Casas, C. Augusto (2001) Tactical asset allocation: An artificial neural network based model. *IJCNN'01. International Joint Conference on Neural Networks. Proceedings (Cat. No.01CH37222)*, Vol. 3 (1), 1811–1816.

Castagneto-Gissey, Giorgio – Nivorozhkin, Eugene (2016) No contagion from Russia toward global equity markets after the 2014 international sanctions. *Economic Analysis and Policy*, Vol. 52 (3), 79–98.

Celebuski, Matthew J. – Hill, Joanne M. – Kilgannon, John J. (1990) Managing currency exposures in international portfolios. *Financial Analysts Journal*, Vol. 46 (1), 16–23.

Ceria, Sebastián – Stubbs, Robert A. (2006) Incorporating estimation errors into portfolio selection: Robust portfolio construction. *Journal of Asset Management*, Vol. 7 (2), 109–127.

Chakraborty, Atreya – Grant, James L. – Trahan, Emery A. (2017) The EVA style approach to tactical asset allocation. *The Journal of Wealth Management*, Vol. 20 (2), 41–53.

Chalabi, Yohan – Würtz, Diethelm (2015) Portfolio allocation. In: *Computational actuarial science with R*, ed. Arthur Charpentier, 447–471. CRC Press, Boca Raton, FL.

Chan, Bertram K. C. (2017) *Applied probabilistic calculus for financial engineering: An introduction using R*. John Wiley & Sons, Hoboken, NJ.

Chan, Ernest P. (2013) *Algorithmic trading: Winning strategies and their rationale*. John Wiley & Sons, Hoboken, NJ.

Chan, Louis K. C. – Jegadeesh, Narasimhan – Lakonishok, Josef (1996) Momentum strategies. *The Journal of Finance*, Vol. 51 (5), 1681–1713.

Chan, Louis K. C. – Jegadeesh, Narasimhan – Lakonishok, Josef (1999) The profitability of momentum strategies. *Financial Analysts Journal*, Vol. 55 (6), 80–90.

Chan, Louis K.C. – Karceski, Jason – Lakonishok, Josef (1999) On portfolio optimization: Forecasting covariances and choosing the risk model. *Review of Financial Studies*, Vol. 12 (5), 937–974.

Chatfield, Chris – Xing, Haipeng (2019) *The analysis of time series: An introduction with R*. 7th ed. Chapman & Hall/CRC Texts in Statistical Science Series. CRC Press, Boca Raton, FL.

Chekhlov, Alexei – Uryasev, Stanislav – Zabarankin, Michael (2000) Portfolio optimization with drawdown constraints. Research Report #2000-5. Center for Applied Optimization and Systems Engineering, University of Florida, Gainesville, FL. <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=223323>, retrieved 1.11.2021.

Chekhlov, Alexei – Uryasev, Stanislav – Zabarankin, Michael (2005) Drawdown measure in portfolio optimization. *International Journal of Theoretical and Applied Finance*, Vol. 8 (1), 13–58.

Chelo, Neil R. (2000) Can volatility be your friend? Journal of Indexes. ETF.com. <https://www.etf.com/publications/journalofindexes/joi-articles/1148.html>, retrieved 1.7.2020.

Chen, Joseph – Hong, Harrison (2002) Discussion of “Momentum and autocorrelation in stock returns”. *The Review of Financial Studies. Special Issue: Conference on Market Frictions and Behavioral Finance*, Vol. 15 (2), 565–573.

Chen, Pierre – Lezmi, Edmond – Roncalli, Thierry – Xu, Jiali (2019) A note on portfolio optimization with quadratic transaction costs. Working Paper. Amundi Asset Management, Paris. <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3683466>, retrieved 15.5.2021.

Chen, Xiaohong – Fan, Yanqin (2006) Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspecification. *Journal of Econometrics*, Vol. 135 (1), 125–154.

Cheng, Minder – Madhavan, Ananth (2009) The dynamics of leveraged and inverse exchange-traded funds. Working Paper. Barclays Global Investors. <https://pdfs.semanticscholar.org/79a3/a012593c9046b511f9cedfdc5a026157ab37.pdf>, retrieved 15.6.2020.

Cherubini, Umberto – Gobbi, Fabio – Mulinacci, Sabrina (2016) *Convolution copula econometrics*. Springer, Cham, Switzerland.

Cherubini, Umberto – Gobbi, Fabio – Mulinacci, Sabrina – Romagnoli, Silvia (2012) *Dynamic copula methods in finance*. John Wiley & Sons, Chichester, West Sussex, UK.

Cheung, Wing (2010) The Black-Litterman model explained. *Journal of Asset Management*, Vol. 11 (4), 229–243.

Chia, Rui Ming Daryl – Lim, Kai Jie Shawn – Chan, Ho Yan Sabrina (2013) Return and risk-return ratio based momentum strategies: A fresh perspective. *Journal of Finance and Investment Analysis*, Vol. 2 (1), 1–13.

Chincarini, Ludwig B. – Kim, Daehwan (2006) *Quantitative equity portfolio management: An active approach to portfolio construction and management*. McGraw-Hill, New York, NY.

Chiodo, Abbigail J. – Owyang, Michael T. (2002) A case study of a currency crisis: The Russian default of 1998. The Federal Reserve Bank of St. Louis Review, Vol. 84 (6), 7–18. <https://files.stlouisfed.org/files/htdocs/publications/review/02/11/ChiodoOwyang.pdf>, retrieved 1.8.2021.

Chiou, Shang C. – Tsay, Ruey S. (2008) A copula-based approach to option pricing and risk assessment. *Journal of Data Science*, Vol. 6 (1), 273–301.

Choie, Kenneth S. – Seff, Eric J. (1989) TIPP: Insurance without complexity: Comment. *The Journal of Portfolio Management*, Vol. 16 (1), 107–108.

Chong, James – Phillips, G. Michael (2014) Tactical asset allocation with macroeconomic factors. *The Journal of Wealth Management*, Vol. 17 (1), 58–69.

Chong, Terence Tai-Leung – Cheng, Sam Ho-Sum – Wong, Elfreda Nga-Yee, (2010) A comparison of stock market efficiency of the BRIC countries. *Technology and Investment*, Vol. 1 (4), 235–238.

Chopra, Vijay K. – Ziemba, William T. (1993) The effect of errors in means, variances, and covariances on optimal portfolio choice. *The Journal of Portfolio Management*, Vol. 19 (2), 6–11.

Chou, Ray Yeutien – Liu, Nathan (2010) The economic value of volatility timing using a range-based volatility model. *Journal of Economic Dynamics & Control*, Vol. 34 (1), 2288–2301.

Choueifaty, Yves – Froidure, Tristan – Reynier, Julien (2013) Properties of the most diversified portfolio. *Journal of Investment Strategies*, Vol. 2 (2), 49–70.

Chow, Tzee-Man – Hsu, Jason C. – Kuo, Li-lan – Li, Feifei (2014) A study of low-volatility portfolio construction methods. *The Journal of Portfolio Management*, Vol. 40 (4), 89–106.

Chu, Ba – Satchell, Stephen (2018) The most entropic canonical copula with an application to 'style' investment. In: *Asymmetric dependence in finance: Diversification, correlation and portfolio management in market downturns*, eds. Jamie Alcock – Stephen Satchell, 221–262. John Wiley & Sons, Chichester, West Sussex.

Claessens, Stijn – Djankov, Simeon – Klingebiel, Daniela (2000) Stock markets in transition economies. *The World Bank Financial Sector Discussion Paper*, Vol. 5 (1), 1–32.

Clarke, Roger – de Silva, Harindra – Thorley, Steven (2002) Portfolio constraints and the fundamental law of active management. *Financial Analysts Journal*, Vol. 58 (5), 48–66.

Clemen, Robert T. – Winkler, Robert L. (1999) Combining probability distributions from experts in risk analysis. *Risk Analysis*, Vol. 19 (2), 187–203.

Clements, A. – Silvennoinen, A. (2013) Volatility timing: How best to forecast portfolio exposures. *Journal of Empirical Finance*, Vol. 24 (1), 108–115.

Compton, William – Kunkel, Robert A. – Kuhlemeyer, Gregory (2013) Calendar anomalies in Russian stocks and bonds. *Managerial Finance*, Vol. 39 (12), 1138–1154.

Connor, Gregory – Korajczyk, Robert A. (1993) A test for the number of factors in an approximate factor model. *The Journal of Finance*, Vol. 48 (4), 1263–1291.

Conrad, Jennifer – Kaul, Gautam (1988) Time-variation in expected returns. *The Journal of Business*, Vol. 61 (4), 409–425.

Cont, Rama (2001) Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance*, Vol. 1 (2), 223–236.

Cont, Rama – Deguest, Romain – Scandolo, Giacomo (2010) Robustness and sensitivity analysis of risk measurement procedures. *Quantitative Finance*, Vol. 10 (6), 593–606.

Cooper, Michael J. – Gutierrez, Roberto C., Jr. – Hameed, Allaudeen (2004) Market states and momentum. *The Journal of Finance*, Vol. 59 (3), 1345–1365.

Cornuéjols, Gérard – Peña, Javier – Tütüncü, Reha (2018) *Optimization methods in finance*. 2nd ed. Cambridge University Press, Cambridge.

Corsi, Fulvio (2009) A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, Vol. 7 (2), 174–196.

Council Regulation (EU) No 960/2014. OJ L 271, 12 September 2014. Council of the European Union. <https://eur-lex.europa.eu/legal-content/GA/TXT/?uri=CELEX:32014R0960>, retrieved 3.6.2018.

Council Regulation (EU) No 1290/2014. OJ L 349, 5 December 2014. Council of the European Union. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32014R1290>, retrieved 3.6.2018.

Coval, Joshua – Stafford, Erik (2007) Asset fire sales (and purchases) in equity markets. *Journal of Financial Economics*, Vol. 86 (1), 479–512.

Creal, Drew – Koopman, Siem Jan – Lucas, André (2012) Generalized autoregressive score models with applications. *Journal of Applied Econometrics*, Vol. 28 (5), 777–795.

Cromwell, Jeff B. – Labys, Walter C. – Terraza, Michel (1994) *Univariate tests for time series models*. Sage University Papers Series: Quantitative Applications in the Social Sciences; No. 99. Sage Publications, Thousand Oaks, CA.

Cvitanic, Jaksa – Karatzas, Ioannis (1998) On dynamic measures of risk. Working Paper. Columbia University, Departments of Mathematics and Statistics, New York, NY. <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=89489>, retrieved 1.4.2021.

Czado, Claudia (2019) *Analyzing dependent data with vine copulas: A practical guide with R*. Springer Nature, Cham, Switzerland.

Czado, Claudia – Brechmann, Eike Christian – Gruber, Lutz (2013) Selection of vine copulas. In: *Copulae in mathematical and quantitative finance: Proceedings of the workshop held Cracow, 10–11 July 2012*, eds. Piotr Jaworski – Fabrizio Durante – Wolfgang Karl Härdle, 17–37. Springer-Verlag, Berlin.

Dai, Min – Zhang, Qing – Zhu, Qiji Jim (2011) Optimal trend following trading rules. Working Paper. National University of Singapore; University of Georgia – Department of Mathematics; Western Michigan University. <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1762118>, retrieved 20.1.2021.

Damodaran, Aswath (2002) *Estimating risk free rates*. Working Paper. Stern School of Business, New York, NY.

Damodaran, Aswath (2008) *What is the riskfree rate? A search for the basic building block*. Working Paper. Stern School of Business, New York, NY.

Damodaran, Aswath (2010) *Into the abyss: What if nothing is risk free?* Working Paper. Stern School of Business, New York, NY.

Daniel, Kent D. – Hirshleifer, David A. – Subrahmanyam, Avanidhar (1997) A theory of overconfidence, self-attribution, and security market under- and over-reactions. Social Science Research Network (SSRN). <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2017>, retrieved 11.6.2019.

Daniel, Kent – Hirshleifer, David – Subrahmanyam, Avanidhar (1998) Investor psychology and security market under- and overreactions. *The Journal of Finance*, Vol 53 (6), 1839–1885.

Daniel, Kent – Moskowitz, Tobias J. (2016) Momentum crashes. *Journal of Financial Economics*, Vol. 122 (1), 221–247.

Danielsson, Jon (2011) *Financial risk forecasting: The theory and practice of forecasting market risk with implementation in R and Matlab*. John Wiley & Sons Ltd, Chichester.

Danielsson, Jon – Jorgensen, Bjorn N. – Samorodnitsky, Gennady – Sarma, Mandira – Vries, Casper G., de (2005) Subadditivity re-examined: The case for Value-at-Risk. Working Paper. Cornell University, New York, NY. <https://people.orie.cornell.edu/gennady/techreports/VaRsubadd.pdf>, retrieved 1.4.2021.

Dasgupta, Amil – Prat, Andrea – Verardo, Michela (2011) The impact of institutional herding. *The Review of Financial Studies*, Vol. 24 (3), 892–925.

Davies, P. L. (1987) Asymptotic behaviour of S-estimates of multivariate location parameters and dispersion matrices. *The Annals of Statistics*, Vol. 15 (3), 1269–1292.

Davies, P. Laurie (1992) The asymptotics of Rousseeuw's minimum volume ellipsoid estimator. *The Annals of Statistics*, Vol 20 (4), 1828–1843.

Davydov, Denis – Vähämaa, Sami (2013) Debt source choices and stock market performance of Russian firms during the financial crisis. *Emerging Markets Review*, Vol. 15 (1), 148–159.

Delcoure, Natalya (2007) The determinants of capital structure in transitional economies. *International Review of Economics & Finance*, Vol. 16 (3), 400–415.

DeLong, J. Bradford – Shleifer, Andrei – Summers, Lawrence H. – Waldmann, Robert J. (1990) Positive feedback investment strategies and destabilizing rational speculation. *The Journal of Finance*, Vol. 45 (2), 379–395.

DeMiguel, Victor – Garlappi, Lorenzo – Uppal, Raman (2009) Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *The Review of Financial Studies*, Vol. 22 (5), 1915–1953.

Demos, Antonis – Sentana, Enrique (1998) An EM algorithm for conditionally heteroscedastic factor models. *Journal of Business & Economic Statistics*, Vol. 16 (3), 357–361.

Dempsey, Mike (2013) The capital asset pricing model (CAPM): The history of a failed revolutionary idea in finance. *A Journal of Accounting, Finance and Business Studies (ABACUS)*, Vol. 49 (1), 7–23.

Dhaene, J. – Laeven, R. J. A. – Vanduffel, S. – Darkiewicz, G. – Goovaerts, M. J. (2008) Can a coherent risk measure be too subadditive? *The Journal of Risk and Insurance*, Vol. 75 (2), 365–386.

Diaz, Antonio – García-Donato, Gonzalo – Mora-Valencia, Andrés (2019) Quantifying risk in traditional energy and sustainable investments. *Sustainability*, Vol. 11 (3), 1–22.

Dichtl, Hubert – Drobetz, Wolfgang (2009) Does tactical asset allocation work? Another look at the fundamental law of active management. *Journal of Asset Management*, Vol. 10 (4), 235–252.

Dichtl, Hubert – Drobetz, Wolfgang (2011) Portfolio insurance and prospect theory investors: Popularity and optimal design of capital protected financial products. *Journal of Banking & Finance*, Vol. 35 (7), 1683–1697.

Dickey, David A. – Fuller, Wayne A. (1979) Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, Vol. 74 (366), 427–431.

Diebold, Francis X. (2015) Comparing predictive accuracy, twenty years later: A personal perspective on the use and abuse of Diebold–Mariano tests. *Journal of Business & Economic Statistics*, Vol. 33 (1), 1–9.

Diebold, Francis X. – Mariano, Roberto S. (1995) Comparing predictive accuracy. *Journal of Business & Economic Statistics*, Vol. 13 (3), 253–263.

Ding, Zhuanxin – Granger, Clive W. J. – Engle, Robert F. (1993) A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, Vol. 1 (1), 83–106.

Disatnik, David J. – Benninga, Simon (2007) Shrinking the covariance matrix: Simpler is better. *The Journal of Portfolio Management*, Vol. 33 (4), 55–63.

Dixit, Avinash K. (1990) *Optimization in economic theory*. 2nd ed. Oxford University Press, New York, NY.

Docherty, Paul – Hurst, Gareth (2018) Return dispersion and conditional momentum returns: International evidence. *Pacific-Basin Finance Journal*, Vol. 50 (1), 263–278.

Donaldson, R. Glen – Kamstra, Mark (1997) An artificial neural network-GARCH model for international stock return volatility. *Journal of Empirical Finance*, Vol. 4 (1), 17–46.

Dong, Ruijun (2009) Differential evolution versus particle swarm optimization for PID controller design. In: *Proceedings of the Fifth International Conference on Natural Computation*, Tianjin, China, August 14–16, 2009, 236–240.

Donoho, David L. – Gasko, Miriam (1992) Breakdown properties of location estimates based on halfspace depth and projected outlyingness. *The Annals of Statistics*, Vol. 20 (4), 1803–1827.

Donohue, Christopher – Yip, Kenneth (2003) Optimal portfolio rebalancing with transaction costs. *The Journal of Portfolio Management*, Vol. 29 (4), 49–63.

Dowd, Kevin (1999) A value at risk approach to risk-return analysis. *The Journal of Portfolio Management*, Vol. 25 (4), 60–67.

Drewes, Sarah (2016) Optimizers everywhere: Optimization in financial applications with MATLAB. MathWorks® Consulting Services. <https://www.mathworks.com/content/dam/mathworks/mathworks-dot-com/company/events/conferences/matlab-computational-finance-conference/2016/proceedings/optimisers-everywhere.pdf >, retrieved 10.3.2021.

Drucker, Harris – Burges, Chris J. C. – Kaufman, Linda – Smola, Alex – Vapnik, Vladimir (1997) Support vector machines. In: *Advances in neural information processing systems 9: Proceedings of the 1996 conference (NIPS-9)*, eds. Michael C. Mozer – Michael I. Jordan – Thomas Petsche, 155–161. The MIT Press, Cambridge, MA.

Duffie, Darrell (1988) *Security markets: Stochastic models*. Economic Theory, Econometrics, and Mathematical Economics. Academic Press, San Diego, CA.

Duffie, Darrell – Pan, Jun (1997) An overview of value at risk. *The Journal of Derivatives*, Vol. 4 (3), 7–49.

Duqi, Andi – Franci, Leonardo – Torluccio, Giuseppe (2014) The Black–Litterman model: The definition of views based on volatility forecasts. *Applied Financial Economics*, Vol. 24 (19), 1285–1296.

Durbin, Michael (2010) *All about high-frequency trading*. McGraw-Hill, New York, NY.

Eichhorn, David – Gupta, Francis – Stubbs, Eric (1998) Using constraints to improve the robustness of asset allocation. *The Journal of Portfolio Management*, Vol. 24 (3), 41–48.

El-Jahel, Lina – Perraudin, William – Sellin, Peter (1999) Value at risk for derivatives. *The Journal of Derivatives*, Vol. 6 (3), 7–26.

Elliott, Graham – Rothenberg, Thomas J. – Stock, James H. (1996) Efficient tests for an autoregressive unit root. *Econometrica*, Vol. 64 (4), 813–836.

Elton, Edwin J. – Gruber, Martin J. (1973) Estimating the dependence structure of share prices-implications for portfolio selection. *The Journal of Finance*, Vol. 28 (5), 1203–1232.

Elton, Edwin J. – Gruber, Martin J. – Brown, Stephen J. – Goetzmann, William N. (2014) *Modern portfolio theory and investment analysis*. 9th ed. John Wiley & Sons, Hoboken, NJ.

Embrechts, Paul – Dias, Alexandra (2004) *Dynamic copula models for multivariate high-frequency data in finance*. Discussion paper. Department of Mathematics, ETH Zurich, Switzerland.

Emmer, Susanne – Kratz, Marie – Tasche, Dirk (2015) What is the best risk measure in practice? A comparison of standard measures. *Journal of Risk*, Vol. 18 (2), 31–60.

Enders, Walter – Liu, Yu (2014) Inflation revisited: New evidence from modified unit root tests. *International Journal of Applied Economics*, Vol. 11 (2), 1–9.

Engle, Robert F. (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, Vol. 50 (1), 987–1008.

Engle, Robert (2002) New frontiers for ARCH models. *Journal of Applied Econometrics*, Vol. 17 (5), 425–446.

Engle, Robert F. – Granger, C. W. J. (1987) Co-integration and error correction: Representation, estimation and testing. *Econometrica*, Vol. 55 (2), 251–276.

Engle, Robert F. – Ng, Victor K. (1993) Measuring and testing the impact of news on volatility. *The Journal of Finance*, Vol. 48 (5), 1749–1778.

Engle, Robert F. – Sheppard, Kevin (2001) Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH. National Bureau of Economic Research (NBER) Working Paper No. 8554. <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=287751>, retrieved 1.8.2021.

Erb, Claude B. – Harvey, Campbell R. (2006) The strategic and tactical value of commodity futures. *Financial Analysts Journal*, Vol. 62 (2), 69–97.

Estrada, Javier (2010) Geometric mean maximization: An overlooked portfolio approach? *The Journal of Investing*, Vol. 19 (4), 134–147.

Fabozzi, Frank J. – Focardi, Sergio – Kolm, Petter N. (2006) *Financial modeling of the equity market: From CAPM to cointegration.* John Wiley & Sons, Hoboken, NJ.

Fabozzi, Frank J. – Focardi, Sergio – Kolm, Petter N. (2010) *Quantitative equity investing: Techniques and strategies.* John Wiley & Sons, Hoboken, NJ.

Fabozzi, Frank J. (2013) *Bond markets, analysis and strategies*. 8th ed. Prentice Hall, Upper Saddle River, NJ.

Fama, Eugene F. (1965) The behaviour of stock market prices. *Journal of Business*, Vol. 38 (1), 34–105.

Fama, Eugene F. (1968) Risk, return and equilibrium: Some clarifying comments. *The Journal of Finance*, Vol. 23 (1), 29–40.

Fama, Eugene F. (1970) Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, Vol. 25 (2), 383–417.

Fama, Eugene F. – French, Kenneth R. (1988) Dividend yields and expected stock returns. *Journal of Financial Economics*, Vol. 22 (1), 3–25.

Fama, Eugene F. – French, Kenneth R. (1989) Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics*, Vol. 25 (1), 23–49.

Fama, Eugene F. – French, Kenneth R. (1992) The cross-section of expected stock returns. *The Journal of Finance*, Vol. 47 (2), 427–465.

Fama, Eugene F. – French, Kenneth R. (2012) Size, value, and momentum in international stock returns. *Journal of Financial Economics*, Vol. 105 (3), 457–472.

Fedorova, E. A. – Pankratov, K. A. (2010) Influence of macroeconomic factors on the Russian stock market. *Studies on Russian Economic Development*, Vol. 21 (2), 165–168.

Fernández, Alberto – Gómez, Sergio (2007) Portfolio selection using neural networks. *Computers & Operations Research*, Vol. 34 (4), 1177–1191.

Figelman, Ilya (2007) Stock return momentum and reversal: A comprehensive study. *The Journal of Portfolio Management*, Vol. 34 (1), 51–69.

Figelman, Ilya (2017) Black–Litterman with a factor structure applied to multi-asset portfolios. *The Journal of Portfolio Management*, Vol. 44 (2), 136–155.

Figlewski, Stephen – Silber, William L. – Subrahmanyam, Marti G. (1990) *Financial options: From theory to practice*. Business One Irwin, Homewood, IL.

Figlewski, Stephen – Wang, Xiaozu (2000) Is the "leverage effect" a leverage effect? Working Paper. NYU Stern School of Business, New York, NY; City University of Hong Kong, Hong Kong. <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=256109>, retrieved 1.4.2021.

Fildes, Robert – Makridakis, Spyros (1995) The impact of empirical accuracy studies on time series analysis and forecasting. *International Statistical Review*, Vol. 63 (3), 289–308.

Filomena, Tiago P. – Lejeune, Miguel A. (2012) Stochastic portfolio optimization with proportional transaction costs: Convex reformulations and computational experiments. *Operations Research Letters*, Vol. 40 (3), 212–217.

Fischer, Matthias – Köck, Christian – Schlüter, Stephan – Weigert, Florian (2009) An empirical analysis of multivariate copula models. *Quantitative Finance*, Vol. 9 (7), 839–854.

Fischer, Thomas – Krauss, Christopher (2018) Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, Vol. 270 (2), 654–669.

Fissler, Tobias – Ziegel, Johanna F. (2016) Higher order elicitability and Osband's principle. *The Annals of Statistics*, Vol. 44 (4), 1680–1707.

Fleming, Jeff – Kirby, Chris – Ostdiek, Barbara (2001) The economic value of volatility timing. *The Journal of Finance*, Vol. 56 (1), 329–352.

Fleming, Jeff – Kirby, Chris – Ostdiek, Barbara (2003) The economic value of volatility timing using "realized" volatility. *Journal of Financial Economics*, Vol. 67 (3), 473–509.

Frahm, Gabriel (2010) Linear statistical inference for global and local minimum variance portfolios. *Statistical Papers*, Vol. 51 (4), 789–812.

Franke, Jürgen – Härdle, Wolfgang Karl – Hafner, Christian Matthias (2015) *Statistics of financial markets: An introduction*. 4th edition. Springer-Verlag, Berlin Heidelberg.

Frazzini, Andrea (2006) The disposition effect and underreaction to news. *The Journal of Finance*, Vol. 61 (4), 2017–2046.

Frazzini, Andrea – Israel, Ronen – Moskowitz, Tobias J. (2012) Trading costs of asset pricing anomalies. Fama-Miller Working Paper. Chicago Booth Research Paper No. 14-05. <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2294498>, retrieved 5.6.2019.

Frazzini, Andrea – Pedersen, Lasse Heje (2014) Betting against beta. *Journal of Financial Economics*, Vol. 111 (1), 1–25.

Freitas, Fabio D. – De Souza, Alberto F. – de Almeida, Ailson R. (2009) Prediction-based portfolio optimization model using neural networks. *Neurocomputing*, Vol. 72 (12), 2155–2170.

Frijns, Bart – Lehnert, Thorsten – Zwinkels, Remco C. J. (2011) Modelling structural changes in the volatility process. *Journal of Empirical Finance*, Vol. 18 (3), 522–532.

Frittelli, Marco – Rosazza Gianin, Emanuela (2002) Putting order in risk measures. *Journal of Banking & Finance*, Vol. 26 (7), 1473–1486.

Frittelli, Marco – Rosazza Gianin, Emanuela (2004) Dynamic convex risk measures. In: *New risk measures for the 21st century*, ed. Giorgio Szegö, 227–248. John Wiley & Sons Ltd, Somerset, NJ.

Frost, Peter A. – Savarino, James E. (1988) For better performance: Constrain portfolio weights. *Journal of Portfolio Management*, Vol. 15 (1), 29–34.

Fusai, Gianluca – Meucci, Attilio (2003) Assessing views. *Risk Magazine*, Vol. 16 (3), 18–21.

Föllmer, Hans – Schied, Alexander (2002) Convex measures of risk and trading constraints. *Finance and Stochastics*, Vol. 6 (4), 429–447.

Gaddy, Clifford G. – Ickes, Barry W. (2010) Russia after the global financial crisis. *Eurasian Geography and Economics*, Vol. 51 (3), 281–311.

Garcia, Rene – Ghysels, Eric (1998) Structural change and asset pricing in emerging markets. *Journal of International Money and Finance*, Vol. 17 (3), 455–473.

Garcia, René – Renault, Éric – Tsafack, Georges (2007) Proper conditioning for coherent VaR in portfolio management. *Management Science*, Vol. 53 (3), 483–494.

Garcia, Sergio (2020) Intermediate to portfolio analysis in R. RPubs by RStudio. <https://rpubs.com/Sergio_Garcia/intermediate_portfolio_analysis_r>, retrieved 25.2.2021.

Garlappi, Lorenzo – Uppal, Raman – Wang, Tan (2006) Portfolio selection with parameter and model uncertainty: A multi-prior approach. *The Review of Financial Studies*, Vol. 20 (1), 41–81.

Garleanu, Nicolae – Pedersen, Lasse Heje (2013) Dynamic trading with predictable returns and transaction costs. *The Journal of Finance*, Vol. 68 (6), 2309–2340.

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

Gavrishchaka, Valeriy V. – Banerjee, Supriya (2006) Support vector machine as an efficient framework for stock market volatility forecasting. *Computational Management Science*, Vol. 3 (2), 147–160.

Geczy, Christopher C. – Samonov, Mikhail (2016) Two centuries of price-return momentum. *Financial Analysts Journal*, Vol. 72 (5), 32–56.

Geidosch, Marco – Fischer, Matthias (2016) Application of vine copulas to credit portfolio risk modeling. *Journal of Risk and Financial Management*, Vol. 9 (4), 1–15.

Genest, Christian – Zidek, James V. (1986) Combining probability distributions: A critique and an annotated bibliography. *Statistical Science*, Vol. 1 (1), 114–135.

George, Thomas J. – Hwang, Chuan-Yang (2004) The 52-week high and momentum investing. *The Journal of Finance*, Vol. 59 (5), 2145–2176.

Geweke, John (1986) Modelling the persistence of conditional variances: A comment. *Econometric Reviews*, Vol. 5 (1), 57–61.

Ghalanos, Alexios (2017) Introduction to the rugarch package: Version 1.3-8. The Comprehensive R Archive Network (CRAN). <https://cran.r-project.org/web/packages/rugarch/vignettes/Introduction_to_the_rugarch_package.pdf>, retrieved 19.7.2019.

Ghalanos, Alexios – Kley, Tobias (2020) Package 'rugarch': Univariate GARCH models. The Comprehensive R Archive Network (CRAN). <https://cran.r-project.org/web/packages/rugarch/rugarch.pdf>, retrieved 14.9.2020.

Ghalanos, Alexios – Rossi, Eduardo – Urga, Giovanni (2015) Independent factor autoregressive conditional density model. *Economic Reviews*, Vol. 34 (5), 594–616.

Ghaoui, Laurent, El – Oks, Maksim – Oustry, Francois (2003) Worst-case Value-at-Risk and robust portfolio optimization: A conic programming approach. *Operations Research*, Vol. 51 (4), 543–556.

Ghosh, Pushpendu – Neufeld, Ariel – Sahoo, Jajati Keshari (2020) Forecasting directional movements of stock prices for intraday trading using LSTM and random forests. Working Paper. Department of Computer Science & Information Systems, BITS Pilani K. K. Birla Goa campus, Goa; Division of Mathematical Sciences, Nanyang Technological University, Singapore; Department of Mathematics, BITS Pilani K. K. Birla Goa campus, Goa. ArXiv. <https://arxiv.org/pdf/2004.10178.pdf>, retrieved 1.6.2020.

Giacometti, Rosella – Bertocchi, Marida – Rachev, Svetlozar T. – Fabozzi, Frank J. (2007) Stable distributions in the Black–Litterman approach to asset allocation. *Quantitative Finance*, Vol. 7 (4), 423–433.

Giorgi, Enrico, De (2002) A note on portfolio selection under various risk measures. Working Paper. Social Science Research Network (SSRN). <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=762104>, retrieved 1.8.2019.

Glosten, Lawrence R. – Jagannathan, Ravi – Runkle, David E. (1993) On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance*, Vol. 48 (5), 1779–1801.

Gnanadesikan, Ramanathan – Kettenring, Jon R. (1972) Robust estimates, residuals and outlier detection with multiresponse data. *Biometrics*, Vol. 28 (1), 81–124.

Gneiting, Tilmann (2011) Making and evaluating point forecasts. *Journal of the American Statistical Association*, Vol. 106 (494), 746–762.

Gochez, Francisco (2015) Notes on the BLCOP package. The Comprehensive R Archive Network (CRAN). <https://cran.r-project.org/web/packages/BLCOP/vignettes/BLCOP.html>, retrieved 14.9.2020.

Gochez, Francisco – Chandler-Mant, Richard – Jin, Suchen – Xie, Jinjing (2020) Package 'BLCOP': Black-Litterman and copula opinion pooling frameworks. The Comprehensive R Archive Network (CRAN). <https://cran.r-project.org/web/packages/BLCOP/BLCOP.pdf>, retrieved 14.9.2020.

Goel, Anubha – Sharma, Amita – Mehra, Aparna (2019) Robust optimization of mixed CVaR STARR ratio using copulas. *Journal of Computational and Applied Mathematics*, Vol. 347 (1), 62–83.

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

Gollinger, Terri L. – Morgan, John B. (1993) Calculation of an efficient frontier for a commercial loan portfolio. *The Journal of Portfolio Management*, Vol. 19 (2), 39–46.

Golub, Bennett W. – Tilman, Leo M. (1997) Measuring yield curve risk using principal components analysis, value at risk, and key rate durations. *The Journal of Portfolio Management*, Vol. 23 (4), 72–84.

Goriaev, Alexei – Sonin, Konstantin (2005) Is political risk company-specific? The market side of the Yukos affair. *Working paper*, New Economic School, Moscow, Russia, 1–32.

Goriaev, Alexei – Zabotkin, Alexei (2006) Risks of investing in the Russian stock market: Lessons of the first decade. *Emerging Markets Review*, Vol. 7 (1), 380–397.

Gorissen, Bram L. – Yanıkoğlu, İhsan – Hertog, Dick, den (2015) A practical guide to robust optimization. *Omega*, Vol. 53 (1), 124–137.

Gourieroux, C. – Laurent, J. P. – Scaillet, O. (2000) Sensitivity analysis of Value at Risk. *Journal of Empirical Finance*, Vol. 7 (4), 225–245.

Gourieroux, C. – Liu, W. (2006) Sensitivity analysis of distortion risk measures. Working Paper. Center for Research in Economics and Statistics (CREST), France; University of Toronto, Canada. <https://ideas.repec.org/p/crs/wpaper/2006-33.html>, retrieved 1.4.2021.

Granger, C. W. J. (1981) Some properties of time series data and their use in econometric model specification. *Journal of Econometrics*, Vol. 16 (1), 121–130.

Grauer, Robert R. – Shen, Frederick C. (2000) Do constraints improve portfolio performance? *Journal of Banking and Finance*, Vol. 24 (1), 1253–1274.

Gray, Wesley R. – Vogel, Jack R. (2016) *Quantitative momentum: A practitioner's guide to building a momentum based stock selection system*. John Wiley & Sons, Hoboken, New Jersey.

Gray, William S. (1993) Historical returns, inflation and future return expectations. *Financial Analysts Journal*, Vol. 49 (4), 35–45.

Green, Richard C. – Hollifield, Burton (1992) When will mean-variance efficient portfolios be well diversified? *The Journal of Finance*, Vol. 47 (5), 1785–1809.

Greff, Klaus – Srivastava, Rupesh Kumar – Koutník, Jan – Steunebrink, Bas R. – Schmidhuber, Jürgen (2017) LSTM: A search space odyssey. *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 28 (10), 2222–2232.

Griffin, John M. – Ji, Xiuqing – Martin, J. Spencer (2003) Momentum investing and business cycle risk: Evidence from pole to pole. *The Journal of Finance*, Vol. 58 (6), 2515–2547.

Grigoriev, Leonid – Salikhov, Marcel (2009) Financial crisis 2008. *Problems of Economic Transition*, Vol. 51 (10), 35–62.

Grinblatt, Mark – Han, Bing (2005) Prospect theory, mental accounting, and momentum. *Journal of Financial Economics*, Vol. 78 (2), 311–339.

Grinblatt, Mark – Keloharju, Matti (2001) What makes investors trade? *The Journal of Finance*, Vol. 56 (2), 589–616.

Grinblatt, Mark – Moskowitz, Tobias J. (2004) Predicting stock price movements from past returns: The role of consistency and tax-loss selling. *Journal of Financial Economics*, Vol. 71 (1), 541–579.

Grinold, Richard C. – Kahn, Ronald N. (2000) *Active portfolio management: A quantitative approach for providing superior returns and controlling risk.* 2nd ed. McGraw-Hill, New York, NY.

Griva, Igor – Nash, Stephen G. – Sofer, Ariela (2009) *Linear and nonlinear optimization*. 2nd ed. Society for Industrial and Applied Mathematics, Philadelphia, PA.

Grundy, Bruce D. – Martin, J. Spencer (2001) Understanding the nature of risks and the source of the rewards to momentum investing. *The Review of Financial Studies*, Vol. 14 (1), 29–78.

Gu, Shihao – Kelly, Bryan – Xiu, Dacheng (2019) Empirical asset pricing via machine learning. Chicago Booth Research Paper No. 18-04; 31st Australasian Finance and Banking Conference 2018; Yale ICF Working Paper No. 2018-09. <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3159577>, retrieved 1.5.2020.

Guidolin, Massimo – Timmermann, Allan (2007) Asset allocation under multivariate regime switching. *Journal of Economic Dynamics & Control*, Vol. 31 (11), 3503–3544.

Gulen, Huseyin – Petkova, Ralitsa (2018) Absolute strength: Exploring momentum in stock returns. Working Paper. Purdue University – Krannert School of Management and Case Western Reserve University – Department of Banking & Finance. <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2638004>, retrieved 1.5.2020.

Gurvich, Evsey – Prilepskiy, Ilya (2015) The impact of financial sanctions on the Russian economy. *Russian Journal of Economics*, Vol. 1 (1), 359–385.

Gutierrez, Roberto C., Jr. – Prinsky, Christo A. (2007) Momentum, reversal, and the trading behaviors of institutions. *Journal of Financial Markets*, Vol. 10 (1), 48–75.

Gyourko, Joseph – Keim, Donald B. (1993) Risk and return in real estate: Evidence from a real estate stock index. *Financial Analysts Journal*, Vol. 49 (5), 39–46.

Györfi, László – Ottucsák, György – Walk, Harro (2012) *Machine learning for financial engineering*. Imperial College Press, London.

Hafner, Christian M. – Manner, Hans (2012) Dynamic stochastic copula models: Estimation, inference and applications. *Journal of Applied Econometrics*, Vol. 27 (2), 269–295.

Hajizadeh, E. – Seifi, A. – Fazel Zarandi, M. H. – Turksen, I. B. (2012) A hybrid modeling approach for forecasting the volatility of S&P 500 index return. *Expert Systems with Applications*, Vol. 39 (1), 431–436.

Hall, Stephen – Urga, Giovanni (2002) Testing for ongoing efficiency in the Russian stock market. Imperial College and City Universe Business School Working Paper. <http://www.e-m-h.org/HaUr02.pdf>, retrieved 4.1.2018.

Halls-Moore, Michael L. (2015) *Successful algorithmic trading*. eBook. QuantStart, United Kingdom.

Halls-Moore, Michael L. (2017) *Advanced algorithmic trading*. eBook. QuantStart, United Kingdom.

Hamilton, James D. (1994) *Time series analysis*. Princeton University Press, Princeton, NJ.

Hamilton, James D. – Susmel, Raul (1994) Autoregressive conditional heteroscedasticity and changes in regime. *Journal of Econometrics*, Vol. 64 (2), 307–333.

Han, Yufeng (2006) Asset allocation with a high dimensional latent factor stochastic volatility model. *The Review of Financial Studies*, Vol. 19 (1), 237–271.

Hansen, Peter Reinhard – Huang, Zhuo – Shek, Howard Howan (2011) Realized GARCH: A joint model for returns and realized measures of volatility. *Journal of Applied Econometrics*, Vol. 27 (6), 877–906.

Hansen, Peter Reinhard – Lunde, Asger (2003) Consistent preordering with an estimated criterion function, with an application to the evaluation and comparison of volatility models. Working Paper, No. 2003-01. Brown University, Department of Economics, Providence, RI; The Aarhus School of Business, Department of Information Science, Aarhus. <https://www.econstor.eu/bitstream/10419/80163/1/363062807.pdf>, retrieved 1.4.2021.

Hansen, Peter R. – Lunde, Asger (2005) A forecast comparison of volatility models: Does anything beat a GARCH(1,1)? *Journal of Applied Econometrics*, Vol. 20 (7), 873–889.

Harlow, W. V. (1991) Asset allocation in a downside-risk framework. *Financial Analysts Journal*, Vol. 47 (5), 28–40.

Harris, Larry (2002) *Trading and exchanges: Market microstructure for practitioners*. Oxford University Press, New York, NY.

Harvey, Campbell R. (1995) Predictable risk and return in emerging markets. *Review of Financial Studies*, Vol. 8 (3), 773–816.

Harvey, Campbell R. – Liechty, John C. – Liechty, Merrill W. – Müller, Peter (2010) Portfolio selection with higher moments. *Quantitative Finance*, Vol. 10 (5), 469–485.

Harvey, David – Leybourne, Stephen – Newbold (1997) Testing the equality of prediction mean squared errors. *International Journal of Forecasting*, Vol. 13 (2), 281–291.

Haslett, John – Raftery, Adrian E. (1989) Space–time modelling with long-memory dependence: Assessing Ireland's wind power resource. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, Vol. 38 (1), 1–50.

Hatherley, Anthony – Alcock, Jamie (2007) Portfolio construction incorporating asymmetric dependence structures: A user's guide. *Accounting and Finance*, Vol. 47 (3), 447–472.

Haugen, Robert A. (2001) *Modern investment theory*. 5th ed. Prentice-Hall, Englewood Cliffs, NJ.

Haugh, Martin B. – Lo, Andrew W. (2001) Asset allocation and derivatives. *Quantitative Finance*, Vol. 1 (1), 45–72.

Hayashi, Fumio (2000) *Econometrics*. Princeton University Press, Princeton, NJ.

Hayo, Bernd – Kutan, Ali M. (2005) The impact of news, oil prices, and global market developments on Russian financial markets. *Economics of Transition*, Vol. 13 (2), 373–393.

He, Guangliang – Litterman, Robert (2002) The intuition behind Black-Litterman model portfolios. Working paper. <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=334304>, retrieved 1.4.2018.

Heaton, J. B. – Polson, N. G. – Witte, J. H. (2016) Deep learning in finance. Working Paper. Booth School of Business, University of Chicago; Department of Mathematics, University College London, and Conjecture LLC. <https://arxiv.org/pdf/1602.06561.pdf>, retrieved 1.5.2020.

Heinen, Andréas – Valdesogo, Alfonso (2009) *Asymmetric CAPM dependence for large dimensions: The canonical vine autoregressive model*. CORE Discussion Papers 2009069. Université catholique de Louvain, Center for Operations Research and Econometrics (CORE), Louvain-la-Neuve, Belgium.

Hellström, Jörgen – Simonsen, Ola (2006) Does the open limit order book reveal information about short-run stock price movements? Working Paper. Umeå Economic Studies 687. Department of Economics, Umeå University, Umeå. <http://www.econ.umu.se/ueslpnr/ues687.pdf>, retrieved 1.6.2020.

Hendry, David F. (1995) *Dynamic econometrics*. Advanced Texts in Econometrics. Oxford University Press, New York, NY.

Henriksson, Roy D. – Merton, Robert C. (1981) On market timing and investment performance. II. Statistical procedures for evaluating forecasting skills. *The Journal of Business*, Vol. 54 (4), 513–533.

Herold, Ulf – Maurer, Raimond (2004) Tactical asset allocation and estimation risk. *Financial Markets and Portfolio Management*, Vol. 18 (1), 39–57.

Herold, Ulf – Maurer, Raimond – Stamos, Michael – Thanh Vo, Huy (2007) Total return strategies for multi-asset portfolios. *The Journal of Portfolio Management*, Vol. 33 (2), 60–76.

Heyde, C. C. – Kou, S. G. – Peng, X. H. (2006) What is a good risk measure: Bridging the gaps between data, coherent risk measures, and insurance risk measures. Working Paper. Columbia University. <https://www.fdic.gov/analysis/cfr/2007/apr/kou.pdf>, retrieved 1.4.2021.

Higgins, M. L. – Bera A. K. (1992) A class of nonlinear ARCH models. *International Economic Review*, Vol. 33 (1), 137–158.

Hofert, Marius – Kojadinovic, Ivan – Mächler, Martin – Yan, Jun (2018) *Elements of copula modeling with R*. Springer International Publishing AG, Cham, Switzerland.

Hoffmann, Manuel – Neuenkirch, Matthias (2017) The pro-Russian conflict and its impact on stock returns in Russia and the Ukraine. *International Economics and Economic Policy*, Vol. 14 (1), 61–73.

Hol, Eugenie – Koopman, Siem Jan (2002) Stock index volatility forecasting with high frequency data. Tinbergen Institute Discussion Paper, TI 2002–068/4. <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=318239>, retrieved 1.4.2021.

Holden, Craig W. – Subrahmanyam, Avanidhar (2002) News events, information acquisition, and serial correlation. *The Journal of Business*, Vol. 75 (1), 1–32.

Hong, Chew Soo – Herk, Leonard F. (1996) Incremental risk aversion and diversification preference. *Journal of Economic Theory*, Vol. 70 (1), 180–200.

Hong, Harrison – Stein, Jeremy C. (1999) A unified theory of underreaction, momentum trading, and overreaction in asset markets. *The Journal of Finance*, Vol. 54 (6), 2143–2184.

Hong, Harrison – Lim, Terence – Stein, Jeremy C. (2000) Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies. *The Journal of Finance*, Vol. 55 (1), 265–295.

Hopper, Gregory P. (1996) Value at risk: A new methodology for measuring portfolio risk. *Business Review (Federal Reserve Bank of Philadelphia)*, July/August 1996, 19–31.

Horváth, Márk – Urbán, András (2010) Growth optimal portfolio selection with short selling and leverage. Draft version. Budapest University of Technology and Economics. <http://www.cs.bme.hu/~oti/portfolio/icpproject/20101212/KuhnTucker.pdf>, retrieved 6.6.2019.

Hsu, Chih-Chiang – Tseng, Chih-Ping – Wang, Yaw-Huei (2008) Dynamic hedging with futures: A copula-based GARCH model. *Journal of Futures Markets*, Vol. 28 (1), 1095–1116.

Huang, Chin-Wen – Hsu, Chun-Pin (2015) Portfolio optimization with GARCH-EVT-copula-CVaR models. *Banking and Finance Review*, Vol. 7 (1), 19–31.

Huber, Peter J. (1964) Robust estimation of a location parameter. *The Annals of Mathematical Statistics*, Vol. 35 (1), 73–101.

Huber, Peter J. (1981) *Robust statistics*. John Wiley & Sons, New York, NY.

Hubert, Mia – Rousseeuw, Peter J. – van Aelst, Stefan (2008) High-breakdown robust multivariate methods. *Statistical Science*, Vol. 23 (1), 92–119.

Hull, John C. (2012) *Risk management and financial institutions*. 3rd ed. Wiley, Hoboken, NJ.

Hull, John C. (2018) *Options, futures, and other derivatives*. 9th ed. Pearson Education Limited, Harlow, Essex.

Hull, John – White, Alan (1987) The pricing of options on assets with stochastic volatilities. *The Journal of Finance*, Vol. 42 (2), 281–300.

Hull, John – White, Alan (1998) Value at risk when daily changes in market variables are not normally distributed. *The Journal of Derivatives*, Vol. 5 (3), 9–19.

Hurst, H. E. (1951) Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, Vol. 116 (1), 770–799.

Hyndman, Rob J. – Athanasopoulos, George (2014) *Forecasting: Principles and practice*. OTexts, Melbourne, Australia.

Hyndman, Rob J. – Athanasopoulos, George – Bergmeir, Christoph – Caceres, Gabriel – Chhay, Leanne – O'Hara-Wild, Mitchell – Petropoulos, Fotios – Razbash, Slava – Wang, Earo – Yasmeen, Farah (2018) Package ‘forecast’: Forecasting functions for time series and linear models. The Comprehensive R Archive Network (CRAN). <https://cran.r-project.org/web/packages/forecast/forecast.pdf>, retrieved 1.4.2018.

Hyndman, Rob J. – Khandakar, Yeasmin (2008) Automatic time series forecasting: The forecast package for R. *Journal of Statistical Software*, Vol. 27 (3), 1–23.

Iakovlev, A. – Danilov, Iu. (1997) The futures market and the development of financial markets in Russia. *Problems of Economic Transition*, Vol. 39 (10), 46–63.

Ibragimov, Rustam (2007) Efficiency of linear estimators under heavy-tailedness: Convolutions of α-symmetric distributions. *Econometric Theory*, Vol. 23 (1), 501–517.

Ibragimov, Rustam (2009) Portfolio diversification and value at risk under thick-tailedness. *Quantitative Finance*, Vol. 9 (5), 565–580.

Ibragimov, Rustam – Walden, Johan (2007) The limits of diversification when losses may be large. *Journal of Banking & Finance*, Vol. 31 (8), 2551–2569.

Idzorek, Thomas M. (2007) A step-by-step guide to the Black-Litterman model: Incorporating user-specified confidence levels. In: *Forecasting expected returns in the financial markets*, ed. Stephen Satchell, 17–38. Academic Press, London.

Ilmanen, Antti (2011) *Expected returns: An investor's guide to harvesting market rewards*. John Wiley & Sons Ltd, Chichester.

Infanger, Gerd (2006) Dynamic asset allocation strategies using a stochastic dynamic programming approach. In: *Handbook of asset and liability management, Volume 1: Theory and methodology*, eds. S. A. Zenios – W. T. Ziemba, 199–251. Elsevier B.V./North-Holland, Amsterdam.

Inkaya, A. – Yolcu Okur, Y. (2014) Analysis of volatility feedback and leverage effects on the ISE30 index using high frequency data. *Journal of Computational and Applied Mathematics*, Vol. 259 (1), 377–384.

Israel, Ronen – Moskowitz, Tobias J. (2013) The role of shorting, firm size, and time on market anomalies. *Journal of Financial Economics*, Vol. 108 (1), 275–301.

Iwan, Mahmud – Akmeliawati, R. – Faisal, Tarig – Al-Assadi, Hayder M. A. A. (2012) Performance comparison of differential evolution and particle swarm optimization in constrained optimization. *Procedia Engineering*, Vol. 41 (1), 1323–1328.

Iwasaki, Ichiro (2014) Global financial crisis, corporate governance, and firm survival: The Russian experience. *Journal of Comparative Economics*, Vol. 42 (1), 178–211.

Jaconetti, Colleen M. – Kinniry, Francis M. Jr. – Zilbering, Yan (2010) Best practices for portfolio rebalancing. Vanguard research. <https://www.van-guard.com/pdf/icrpr.pdf>, retrieved 25.9.2017.

Jagannathan, Ravi – Ma, Tongshu (2003) Risk reduction in large portfolios: Why imposing the wrong constraints helps. *Journal of Finance*, Vol 58 (1), 1651–1683.

Jain, Prayut – Jain, Shashi (2019) Can machine learning-based portfolios outperform traditional risk-based portfolios? The need to account for covariance misspecification. *Risks*, Vol. 7 (3):74, 1–27.

Jalolov, Mirzosharif – Miyakoshi, Tatsuyoshi (2005) Who drives the Russian financial markets? *The Developing Economies*, Vol. 43 (3), 374–395.

Jarque, Carlos M. – Bera, Anil K. (1980) Efficient tests for normality, homoskedasticity and serial independence of regression residuals. *Economics Letters*, Vol. 6 (3), 255–259.

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

Jegadeesh, Narasimhan – Titman, Sheridan (1995a) Short-horizon return reversals and the bid-ask spread. *Journal of Financial Intermediation*, Vol. 4 (2), 116–132.

Jegadeesh, Narasimhan – Titman, Sheridan (1995b) Overreaction, delayed reaction, and contrarian profits. *Review of Financial Studies*, Vol. 8 (4), 973–993.

Jensen, Michael C. (1968) The performance of mutual funds in the period 1945–1964. *The Journal of Finance*, Vol. 23 (2), 389–416.

Jin, Hanqing – Markowitz, Harry – Yu Zhou, Xun (2006) A note on semivariance. *Mathematical Finance*, Vol. 16 (1), 53–61.

Jing-rong, Dong – Yu-Ke, Chen – Yan, Zou (2011) Combining stock market volatility forecasts with analysis of stock materials under regime switching. In: *Advances in Computer Science, Intelligent System and Environment*, eds. David Jin – Sally Lin, 393–402. Advances in Intelligent and Soft Computing, Vol. 104. Springer, Berlin, Heidelberg.

Jobson, J. D. – Korkie, Bob M. (1981) Putting Markowitz theory to work. *Journal of Portfolio Management*, Vol. 7 (1), 70–74.

Johannes, Michael – Polson, Nicholas – Stroud, Jon (2002) Sequential optimal portfolio performance: Market and volatility timing. Working Paper. The Graduate School of Business, Columbia University, NY; The Graduate School of Business, University of Chicago, IL; The Wharton School, University of Pennsylvania, PA. <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=304976>, 1.7.2021.

Johansen, Søren (1995) *Likelihood-based inference in cointegrated vector autoregressive models*. Advanced Texts in Econometrics. Oxford University Press, Oxford.

Johansen, Søren – Juselius, Katarina (1990) Maximum likelihood estimation and inference on cointegration – with applications to the demand for money. *Oxford Bulletin of Economics and Statistics*, Vol. 52 (2), 169–210.

Johansen, Søren – Juselius, Katarina (1992) Testing structural hypotheses in a multivariate cointegration analysis of the PPP and the UIP for UK. *Journal of Econometrics*, Vol. 53 (3), 211–244.

Johansson, Frederik – Seiler, Michael J. – Tjarnberg, Mikael (1999) Measuring downside portfolio risk. *The Journal of Portfolio Management*, Vol. 26 (1), 96–107.

Johnson, Norman L. – Kotz, Samuel – Balakrishnan, N. (1994) *Continuous univariate distributions: Volume 1*. 2nd ed. Distributions in Statistics Series. John Wiley & Sons, New York, NY.

Jondeau, Eric – Jurczenko, Emmanuel – Rockinger, Michael (2018) Moment component analysis: An illustration with international stock markets. *Journal of Business & Economic Statistics*, Vol. 36 (4), 576–598.

Jondeau, Eric – Rockinger, Michael (2006a) Optimal portfolio allocation under higher moments. *European Financial Management*, Vol. 12 (1), 29–55.

Jondeau, Eric – Rockinger, Michael (2006b) The copula-GARCH model of conditional dependencies: An international stock market application. *Journal of International Money and Finance*, Vol. 25 (5), 827–853.

Jondeau, Eric – Rockinger, Michael (2012) On the importance of time variability in higher moments for asset allocation. *Journal of Financial Econometrics*, Vol. 10 (1), 84–123.

Jones, Robert – Lim, Terence – Zangari, Peter J. (2007) The Black-Litterman model for structured equity portfolios: Using views based on factors to construct equity portfolios. *The Journal of Portfolio Management*, Vol. 33 (2), 24–33.

Jorion, Philippe (1985) International portfolio diversification with estimation risk. *Journal of Business*, Vol. 58 (1), 259–278.

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

Jorion, Philippe (1997) Lessons from the Orange County bankruptcy. *The Journal of Derivatives*, Vol. 4 (4), 61–66.

Jurczenko, Emmanuel (2015) *Risk-based and factor investing*. 1st ed. ISTE Press – Elsevier, London.

Jurczenko, Emmanuel – Teiletche, Jérôme (2015) Risk-based investing but what risk(s)? In: *Risk-based and factor investing*, ed. Emmanuel Jurczenko, 147–171. ISTE Press – Elsevier, London.

Kahneman, Daniel – Riepe, Mark W. (1998) Aspects of investor psychology. *The Journal of Portfolio Management*, Vol. 24 (4), 52–65.

Kahneman, Daniel – Tversky, Amos (1979) Prospect theory: An analysis of decision under risk. *Econometrica*, Vol. 47 (2), 263–291.

Kan, Raymond – Zhou, Guofu (2007) Optimal portfolio choice with parameter uncertainty. *The Journal of Financial and Quantitative Analysis*, Vol. 42 (3), 621–656.

Kandel, Shmuel – Stambaugh, Robert F. (1996) On the predictability of stock returns: An asset-allocation perspective. *The Journal of Finance*, Vol. 51 (2), 385–424.

Karas, Alexei – Schoors, Koen J. L. – Weill, Laurent (2008) Are private banks more efficient than public banks? Evidence from Russia. BOFIT Discussion Paper No. 3/2008. <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1121709>, retrieved 15.5.2019.

Katz, Erez (2019) Measure your portfolio's statistical significance beyond Sharpe ratio. Blog Post. Lucena Research: Predictive Analytics. <https://lucenaresearch.com/2019/01/05/portfolio-statistically-significant/>, retrieved 1.4.2021.

Kaya, Hakan – Lee, Wai (2012) Demystifying risk parity. Working Paper. Social Science Research Network (SSRN). <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1987770&download=yes>, retrieved 1.8.2021.

Kelly, J. L., Jr. (1956) A new interpretation of information rate. *The Bell System Technical Journal*, Vol. 35 (4), 917–926.

Kempf, Alexander – Memmel, Christoph (2006) Estimating the global minimum variance portfolio. *Schmalenbach Business Review*, Vol. 58 (4), 332–348.

Kijima, Masaaki (1997) The generalized harmonic mean and a portfolio problem with dependent assets. *Theory and Decision*, Vol. 43 (1), 71–87.

Kim, Chang-Jin – Morley, James C. – Nelson, Charles R. (2001) Does an intertemporal tradeoff between risk and return explain mean reversion in stock prices? *Journal of Empirical Finance*, Vol. 8 (4), 403–426.

Kim, Dongcheol – Kon, Stanley J. (1994) Alternative models for the conditional heteroscedasticity of stock returns. *The Journal of Business*, Vol. 67 (4), 563–598.

Kim, Saejoon (2017) Volatility forecasting for low-volatility portfolio selection in the US and the Korean equity markets. *Journal of Experimental & Theoretical Artificial Intelligence*, Vol. 30 (1), 71–88.

Kinlaw, Will – Kritzman, Mark (2009) Optimal currency hedging in- and out-of-sample. *Journal of Asset Management*, Vol. 10 (1), 22–36.

Kinnunen, Jyri (2013) Dynamic return predictability in the Russian stock market. *Emerging Markets Review*, Vol. 15 (1), 107–121.

Kinnunen, Jyri – Martikainen, Minna (2017) Expected returns and idiosyncratic risk: Industry-level evidence from Russia. *Emerging Markets Finance and Trade*, Vol. 53 (11), 2528–2544.

Kipnis, Ilya (2016) How to compute turnover with Return.Portfolio in R. Blog Post. R-bloggers. <https://www.r-bloggers.com/2016/05/how-to-compute-turnover-with-return-portfolio-in-r/>, retrieved 17.5.2021.

Klinker, Frank (2011) Exponential moving average versus moving exponential average. *Mathematische Semesterberichte*, Vol. 58 (1), 97–107.

Koesterich, Russ (2014) Must-know: Why Indian equities seem overvalued. Yahoo Finance. <https://finance.yahoo.com/news/must-know-why-indian-equities-220105275.html>, retrieved 25.9.2018.

Koijen, Ralph S. J. – Moskowitz, Tobias J. – Pedersen, Lasse Heje – Vrugt, Evert B. (2018) Carry. *Journal of Financial Economics*, Vol. 127 (2), 197–225.

Kolm, Petter N. – Ritter, Gordon (2015) Multiperiod portfolio selection and Bayesian dynamic models. *Risk*, Vol. 28 (3), 50–54.

Kolm, Petter N. – Tütüncü, Reha – Fabozzi, Frank (2014) 60 years of portfolio optimization: Practical challenges and current trends. *European Journal of Operational Research*, Vol. 234 (2), 356–371.

Konno, Hiroshi – Wijayanayake, Annista (1999) Mean-absolute deviation portfolio optimization model under transaction costs. *Journal of the Operations Research Society of Japan*, Vol. 42 (4), 422–435.

Korajczyk, Robert A. – Sadka, Ronnie (2004) Are momentum profits robust to trading costs? *The Journal of Finance*, Vol. 59 (3), 1039–1082.

Kou, Steven – Peng, Xianhua – Heyde, Chris C. (2013) External risk measures and Basel Accords. *Mathematics of Operations Research*, Vol. 38 (3), 393–417.

Kratz, Marie – Lok, Yen H. – McNeil, Alexander J. (2018) Multinomial VaR backtests: A simple implicit approach to backtesting expected shortfall. *Journal of Banking and Finance*, Vol. 88 (1), 393–407.

Krauss, Christopher – Do, Xuan Anh – Huck, Nicolas (2016) Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500. Working paper. FAU Discussion Papers in Economics, No. 03/2016. Friedrich-Alexander-Universität Erlangen-Nürnberg, Institute for Economics, Erlangen. <https://www.econstor.eu/bitstream/10419/130166/1/856307327.pdf>, retrieved 22.5.2020.

Krishnan, Hari – Mains, Norman (2005) The two-factor Black-Litterman model. Risk Magazine. <https://www.risk.net/risk-management/1500309/two-factor-black-litterman-model>, retrieved 23.5.2019.

Kritzman, Mark (1986) What's wrong with portfolio insurance? *The Journal of Portfolio Management*, Vol. 13 (1), 13–16.

Kritzman, Mark (1990) Portfolio insurance and option replication strategies. In: *Financial options: From theory to practice*, eds. Stephen Figlewski – William L. Silber – Marti G. Subrahmanyam, 454–484. Business One Irwin, Homewood, IL.

Kritzman, Mark – Li, Yuanzhen – Page, Sébastien – Rigobon, Roberto (2011) Principal components as a measure of systematic risk. *The Journal of Portfolio Management*, Vol. 37 (4), 112–126.

Kritzman, Mark – Page, Sébastien – Turkington, David (2010) In defense of optimization: The fallacy of 1/N. *Financial Analysts Journal*, Vol. 66 (2), 31–39.

Kritzman, Mark – Page, Sébastien – Turkington, David (2012) Regime shifts: Implications for dynamic strategies. *Financial Analysts Journal*, Vol. 68 (3), 22–39.

Krokhmal, Pavlo A. (2007) Higher moment coherent risk measures. *Quantitative Finance*, Vol. 7 (4), 373–387.

Krokhmal, Pavlo – Palmquist, Jonas – Uryasev, Stanislav (2001) Portfolio optimization with conditional Value-at-Risk objective and constraints. Working Paper. University of Florida, Department of Industrial and Systems Engineering, Gainesville, FL; McKinsey & Company, Stockholm, Sweden. <https://www.ise.ufl.edu/uryasev/files/2011/11/kro_CVaR.pdf>, retrieved 1.5.2021.

Krokhmal, Pavlo – Uryasev, Stanislav – Zrazhevsky, Grigory (2002) Risk management for hedge fund portfolios. *The Journal of Alternative Investments*, Vol. 5 (1), 10–29.

Krokhmal, Pavlo – Uryasev, Stanislav – Zrazhevsky, Grigory (2005) Numerical comparison of Conditional Value-at-Risk and Conditional Drawdown-at-Risk approaches: Application to hedge funds. In: *Applications of stochastic programming*, eds. Stein W. Wallace – William T. Ziemba, 609–631. MPS-SIAM Series on Optimization. The Society for Industrial and Applied Mathematics and the Mathematical Programming Society, Philadelphia, PA.

Kudrin, Alexei L. (2009) The global financial crisis and its effect on Russia. *Problems of Economic Transition*, Vol. 52 (3), 4–29.

Kudrin, Alexey – Gurvich, Evsey (2015) A new growth model for the Russian economy. *Russian Journal of Economics*, Vol. 1 (1), 30–54.

Kudrov, Alexander – Zlotnik, Andrew – Dukhovnaya, Elena – Fantazzini, Dean (2012) Short selling in Russia: Main regulations and empirical evidence from medium- and long-term portfolio strategies. In: *Handbook of short selling*, ed. Greg N. Gregoriou, 387–400. Elsevier, Waltham, MA.

Kupiec, Paul H. (1998) Stress testing in a value at risk framework. *The Journal of Derivatives*, Vol. 6 (1), 7–24.

Kusuoka, Shigeo (2001) On law invariant coherent risk measures. *Advances in Mathematical Economics*, Vol. 3 (1), 83–95.

Larsen, Glen A., Jr. – Resnick, Bruce G. (2001) Parameter estimation techniques, optimization frequency, and portfolio return enhancement. *The Journal of Portfolio Management*, Vol. 27 (4), 27–34.

Lassance, Nathan – DeMiguel, Victor – Vrins, Frédéric (2018) Optimal portfolio diversification via independent component analysis. Working Paper. Social Science Research Network (SSRN). <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3285156&download=yes>, retrieved 1.8.2021.

Lazo, Juan G. Lazo – Pacheco, Marco Aurélio C. – Vellasco, Marley Maria R. (2002) Portfolio selection and management using a hybrid intelligent and statistical system. In: *Genetic algorithms and genetic programming in computational finance*, ed. Shu-Heng Chen, 221–238. Springer Science+Business Media, New York, NY.

LeDell, Erin – Gill, Navdeep – Aiello, Spencer – Fu, Anqi – Candel, Arno – Click, Cliff – Kraljevic, Tom – Nykodym, Tomas – Aboyoun, Patrick – Kurka, Michal – Malohlava, Michal – Rehak, Ludi – Eckstrand, Eric – Hill, Brandon – Vidrio, Sebastian – Jadhawani, Surekha – Wang, Amy – Peck, Raymond – Wong, Wendy – Gorecki, Jan – Dowle, Matt – Tang, Yuan – DiPerna, Lauren – H2O.ai (2020) Package 'h2o': R interface for the 'H2O' scalable machine learning platform. The Comprehensive R Archive Network (CRAN). <https://cran.r-project.org/web/packages/h2o/h2o.pdf>, retrieved 19.10.2020.

Ledoit, Olivier – Wolf, Michael (2003) Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of Empirical Finance*, Vol. 10 (5), 603–621.

Ledoit, Olivier – Wolf, Michael (2004) Honey, I shrunk the sample covariance matrix. *Journal of Portfolio Management*, Vol. 30 (1), 110–117.

Lee, Brian – Rowe, Yung (2016) Package 'tawny': Clean covariance matrices using random matrix theory and shrinkage estimators for portfolio optimization. The Comprehensive R Archive Network (CRAN). <https://cran.r-project.org/web/packages/tawny/index.html>, retrieved 23.2.2021.

Lee, Charles M. C. – Swaminathan, Bhaskaran (2000) Price momentum and trading volume. *The Journal of Finance*, Vol. 55 (5), 2017–2069.

Lee, Wai (1998) Return and risk characteristics of tactical asset allocation under imperfect information. *The Journal of Portfolio Management*, Vol. 25 (1), 61–70.

Lee, Wai (2000a) Advanced theory and methodology of tactical asset allocation. Unpublished Manuscript. Duke's Fuqua School of Business, Durham, NC. <https://faculty.fuqua.duke.edu/~charvey/Teaching/IntesaBci_2001/Leex.pdf>, retrieved 1.8.2021.

Lee, Wai (2000b) *Theory and methodology of tactical asset allocation*. 1st edition. Wiley & Sons, New Hope, Pennsylvania.

Lehar, Alfred – Scheicher, Martin – Schittenkopf, Christian (2002) GARCH vs. stochastic volatility: Option pricing and risk management. *Journal of Banking & Finance*, Vol. 26 (3), 323–345.

Lehmann, Bruce N. (1990) Fads, martingales, and market efficiency. *The Quarterly Journal of Economics*, Vol. 105 (1), 1–28.

Lejeune, Miguel A. (2011) A VaR Black–Litterman model for the construction of absolute return fund-of-funds. *Quantitative Finance*, Vol. 11 (10), 1489–1501.

Leland, Hayne E. (1999) Beyond mean–variance: Performance measurement in a non-symmetrical world (corrected). *Financial Analysts Journal*, Vol. 55 (1), 27–36.

Leland, Hayne – Rubinstein, Mark (1976) The evolution of portfolio insurance. In: *Portfolio insurance: A guide to dynamic hedging*, eds. Donald Luskin, chapter 2. John Wiley & Sons, Hoboken, NJ.

Lesmond, David A. – Schill, Michael J. – Zhou, Chunsheng (2004) The illusory nature of momentum profits. *Journal of Financial Economics*, Vol. 71 (1), 349–380.

Lewis, Nigel – Okunev, John – White, Derek (2007) Using a Value at Risk approach to enhance tactical asset allocation. *The Journal of Investing*, Vol. 16 (4), 100–107.

Li, Bin – Hoi, Steven C. H. (2012) On-line portfolio selection with moving average reversion. Proceedings of the 29th International Conference on Machine Learning. <https://arxiv.org/ftp/arxiv/papers/1206/1206.4626.pdf>, retrieved 3.7.2019.

Li, Bin – Hoi, Steven C. H. (2014) Online portfolio selection: A survey. *ACM Computing Surveys*, Vol. 46 (3), 1–33.

Li, Bin – Hoi, Steven C. H. (2016) *Online portfolio selection: Principles and algorithms*. CRC Press, Boca Raton, FL.

Li, Wei – Lam, Kin (2002) Optimal market timing strategies under transaction costs. *Omega*, Vol. 30 (1), 97–108.

Liljeblom, Eva – Maury, Benjamin (2016) Shareholder protection, ownership, and dividends: Russian evidence. *Emerging Markets Finance & Trade*, Vol. 52 (1), 2414–2433.

Linsmeier, Thomas J. – Pearson, Neil D. (2000) Value at risk. *Financial Analysts Journal*, Vol. 56 (2), 47–67.

Lintner, John (1965) The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics*, Vol. 47 (1), 13–37.

Litterman, Bob (2003) Beyond equilibrium, the Black-Litterman approach. In: *Modern investment management: An equilibrium approach*, eds. Bob Litterman and the Quantitative Resources Group, Goldman Sachs Asset Management, 76–88. John Wiley & Sons, Hoboken, NJ.

Litterman, Robert (1996) Hot spots™ and hedges. *The Journal of Portfolio Management*, Vol. 23 (5), 52–75.

Litterman, Robert – Winkelmann, Kurt (1998) Estimating covariance matrices. Research Paper. Risk Management Series, Goldman Sachs. <https://faculty.fuqua.duke.edu/~charvey/Teaching/IntesaBci_2001/GS_Estimating_covariance_matrices.pdf>, retrieved 1.8.2021.

Little, Roderick J. A. – Rubin, Donald B. (2002) *Statistical analysis with missing data*. 2nd ed. John Wiley & Sons, Hoboken, NJ.

Liu, Hong – Loewenstein, Mark (2002) Optimal portfolio selection with transaction costs and finite horizons. *The Review of Financial Studies*, Vol. 15 (3), 805–835.

Liu, Jun – Longstaff, Francis A. (2004) Losing money on arbitrage: Optimal dynamic portfolio choice in markets with arbitrage opportunities. *The Review of Financial Studies*, Vol. 17 (3), 611–641.

Liu, Jun – Pan, Jun (2003) Dynamic derivative strategies. *Journal of Financial Economics*, Vol. 69 (3), 401–430.

Ljung, Greta M. – Box, George E. P. (1978) On a measure of lack of fit in time series models. *Biometrica*, Vol. 65 (2), 297–303.

Lo, Andrew W. (2004) The adaptive markets hypothesis: Market efficiency from an evolutionary perspective. *Journal of Portfolio Management*, Vol. 30 (5), 15–29.

Lo, Andrew W. – MacKinlay, A. Craig (1988) Stock market prices do not follow random walks: Evidence from a simple specification test. *The Review of Financial Studies*, Vol. 1 (1), 41–66.

Lo, Andrew W. – MacKinlay, A. Craig (1990) When are contrarian profits due to stock market overreaction? *The Review of Financial Studies*, Vol. 3 (2), 175–205.

Lo, Andrew W. – MacKinlay (1997) Maximizing predictability in the stock and bond markets. *Macroeconomic Dynamics*, Vol. 1 (1), 102–134.

Lobo, Miguel Sousa (2000) *Robust and convex optimization with applications in finance*. Doctoral dissertation. Stanford University, Stanford, CA.

Lobo, Miguel Sousa – Boyd, Stephen (2000) The worst-case risk of a portfolio. Unpublished Technical Report. Duke University; Stanford University. <https://web.stanford.edu/~boyd/papers/pdf/risk_bnd.pdf>, retrieved 4.3.2021.

Lobo, Miguel Sousa – Vandenberghe, Lieven – Boyd, Stephen – Lebret, Hervé (1998) Applications of second-order cone programming. *Linear Algebra and its Applications*, Vol. 284 (1), 193–228.

Logoveev, Alexander – Cherinko, Gregory (2015) Recent periods of financial turbulence on the Russian stock market and their effect on price correlation and value at risk. *Undergraduate Economic Review*, Vol. 11 (1), 1–10.

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

López de Prado, Marcos (2018) *Advances in financial machine learning*. John Wiley & Sons, Hoboken, New Jersey.

Lopuhaä, Hendrik P. (1989) On the relation between S-estimators and M-estimators of multivariate location and covariance. *The Annals of Statistics*, Vol. 17 (4), 1662–1683.

Lopuhaä, Hendrik P. (1991) Multivariate τ-estimators for location and scatter. *The Canadian Journal of Statistics*, Vol. 19 (3), 307–321.

Lopuhaä, Hendrik P. (1992) Highly efficient estimators of multivariate location with high breakdown point. *The Annals of Statistics*, Vol. 20 (1), 398–413.

Lou, Dong (2012) A flow-based explanation for return predictability. *The Review of Financial Studies*, Vol. 25 (12), 3457–3489.

Low, Rand Kwong Yew – Alcock, Jamie – Faff, Robert – Brailsford, Timothy (2018) Canonical vine copulas in the context of modern portfolio management: Are they worth it? In: *Asymmetric dependence in finance: Diversification, correlation and portfolio management in market downturns*, eds. Jamie Alcock – Stephen Satchell, 263–289. John Wiley & Sons, Chichester, West Sussex.

Lozinskaia, Agata M. – Saltykova, Anastasiia D. (2019) Fundamental factors affecting the MOEX Russia index: Structural break detection in a long-term time series. Working paper. HSE Research Paper No. WP BRP 77/FE/2019, National Research University Higher School of Economics. <https://wp.hse.ru/data/2019/10/16/1527654019/77FE2019.pdf>, retrieved 1.8.2021.

Lucey, Brian M. – Voronkova, Svitlana (2005) *Russian equity market linkages before and after the 1998 crisis: Evidence from time-varying and stochastic cointegration tests*. BOFIT Discussion Papers 12/2005, Helsinki.

Lucey, Brian M. – Voronkova, Svitlana (2008) Russian equity market linkages before and after the 1998 crisis: Evidence from stochastic and regime-switching cointegration tests. *Journal of International Money and Finance*, Vol. 27 (1), 1303–1324.

Luciano, Elisa – Schoutens, Wim (2006) A multivariate jump-driven financial asset model. *Quantitative Finance*, Vol. 6 (5), 385–402.

Luenberger, David G. (1998) *Investment science*. Oxford University Press, New York, NY.

Lustig, Yoram (2013) *Multi-asset investing: A practical guide to modern portfolio management*. Harriman House, Hampshire.

Lütkepohl, Helmut (2004) Vector autoregressive and vector error correction models. In: *Applied time series econometrics*, eds. Helmut Lütkepohl – Markus Krätzig, 86–158. Cambridge University Press, Cambridge.

Lütkepohl, Helmut (2005) *New introduction to multiple time series analysis*. Springer-Verlag, Berlin, Heidelberg.

Luukka, Pasi – Pätäri, Eero – Fedorova, Elena – Garanina, Tatiana (2016) Performance of moving average trading rules in a volatile stock market: The Russian evidence. *Emerging Markets Finance and Trade*, Vol. 52 (10), 2434–2450.

Lyons, Richard K. (2001) *The microstructure approach to exchange rates*. The MIT Press, Cambridge, MA.

Lyuu, Yuh-Dauh (2004) *Financial engineering and computation: Principles, mathematics, algorithms*. Cambridge University Press, Cambridge.

MacBeth, James D. – Emanuel, David C. (1993) Tactical asset allocation: Pros and cons. *Financial Analysts Journal*, Vol. 49 (6), 30–43.

MacIntyre, Stacy – Hess, Tim (2019) EIA forecasts world crude oil prices to rise gradually, averaging $65 per barrel in 2020. U.S. Energy Information Administration (EIA). <https://www.eia.gov/todayinenergy/detail.php?id=38032>, retrieved 13.5.2019.

MacKenzie, Donald – Spears, Taylor (2014) 'The formula that killed Wall Street': The Gaussian copula and modelling practices in investment banking. *Social Studies of Science*, Vol. 44 (3), 393–417.

MacMinn, Richard D. (1984) A general diversification theorem: A note. *The Journal of Finance*, Vol. 39 (2), 541–550.

Maillard, Sébastien – Roncalli, Thierry – Teïletche, Jérôme (2010) The properties of equally weighted risk contribution portfolios. *The Journal of Portfolio Management*, Vol. 36 (4), 60–70.

Malkiel, Burton G. (2019) *A random walk down Wall Street: The time-tested strategy for successful investing*. 12th rev. ed. W. W. Norton & Company, New York, NY.

Malkiel, Burton G. – Xu, Yexiao (1997) Risk and return revisited. *The Journal of Portfolio Management*, Vol. 23 (3), 9–14.

Mansini, Renata – Ogryczak, Włodzimierz – Speranza, M. Gracia (2015) *Linear and mixed integer programming for portfolio optimization*. EURO Advanced Tutorials on Operational Research. Springer International Publishing, Cham.

Markowitz, Harry M. (1952) Portfolio selection. *The Journal of Finance*, Vol. 7 (1), 77–91.

Markowitz, Harry M. (1959) *Portfolio selection: efficient diversification of investments*. John Wiley & Sons, New York.

Markowitz, Harry M. (1987) *Mean-variance analysis in portfolio choice and capital markets*. Basil Blackwell, Oxford, UK.

Markowitz, Harry M. (1994) The general mean-variance portfolio selection problem. *Philosophical Transactions of the Royal Society of London. Series A: Physical and Engineering Sciences*, Vol. 347 (1684), 543–549.

Markowitz, Harry M. (1999) The early history of portfolio theory: 1600–1960. *Financial Analysts Journal*, Vol. 55 (4), 5–16.

Markowitz, Harry M. – van Dijk, Erik (2006) Risk-return analysis. In: *Handbook of asset and liability management, Volume 1: Theory and methodology*, eds. S. A. Zenios – W. T. Ziemba, 139–197. Elsevier B.V./North-Holland, Amsterdam.

Maronna, Ricardo A. – Martin, R. Douglas – Yohai, Victor J. – Salibián-Barrera, Matías (2019) *Robust statistics: Theory and methods (with R)*. 2nd ed. John Wiley & Sons Ltd, Hoboken, NJ.

Maronna, Ricardo A. – Zamar, Ruben H. (2002) Robust estimates of location and dispersion of high-dimensional datasets. *Technometrics*, Vol. 44 (4), 307–317.

Marquering, Wessel – Verbeek, Marno (2004) The economic value of predicting stock index returns and volatility. *Journal of Financial and Quantitative Analysis*, Vol. 39 (2), 407–429.

Marshall, John F. – Bansal, Vipul K. (1992) *Financial engineering: A complete guide to financial innovation*. New York Institute of Finance, New York, NY.

Martellini, Lionel – Ziemann, Volker (2007) Extending Black-Litterman analysis beyond the mean-variance framework. *The Journal of Portfolio Management*, Vol. 33 (4), 33–44.

Martellini, Lionel – Ziemann, Volker (2008) Improved forecasts of higher-order co-moments and implications for portfolio selection. EDHEC Risk and Asset Management Research Centre. <http://docs.edhec-risk.com/EAID-2008-Doc/documents/Higher_Order_Comoments.pdf>, retrieved 18.10.2017.

Martellini, Lionel – Ziemann, Volker (2009) Improved estimates of higher-order comoments and implications for portfolio selection. *The Review of Financial Studies*, Vol. 23 (4), 1467–1502.

Martens, Martin (2002) Measuring and forecasting S&P 500 index-futures volatility using high-frequency data. *The Journal of Futures Markets*, Vol. 22 (6), 497–518.

Martin, R. Douglas – Rachev, Svetlozar Zari – Siboulet, Frederic (2003) Phi-alpha optimal portfolios and extreme risk management. *Wilmott Magazine of Finance (November)*, 70–83.

Maslov, Sergei – Mills, Mark (2001) Price fluctuations from the order book perspective: Empirical facts and a simple model. *Physica A*, Vol. 299 (2), 234–246.

Mason, Scott P. – Merton, Robert C. – Perold, André F. – Tufano, Peter (1995) *Cases in financial engineering: Applied studies of financial innovation*. Prentice Hall, Englewood Cliffs, NJ.

MATLAB (2020) Overview: Math, graphics, programming. MathWorks®. <https://www.mathworks.com/products/matlab.html>, retrieved 1.5.2020.

Mau, Vladimir (2017) Lessons in stabilization and prospects for growth: Russia's economic policy in 2016. *Russian Journal of Economics*, Vol. 3 (1), 109–128.

Mau, Vladimir – Ulyukaev, Alexey (2015) Global crisis and challenges for Russian economic development. *Russian Journal of Economics*, Vol. 1 (1), 4–29.

Mausser, Helmut – Rosen, Dan (2000) Managing risk with expected shortfall. In: *Probabilistic constrained optimization: Methodology and applications*, ed. Stanislav P. Uryasev, 198–219. Springer Science+Business Media, Dordrecht.

McAllister, Ian – White, Stephen (2011) Democratization in Russia and the global financial crisis. *Journal of Communist Studies and Transition Politics*, Vol. 27 (3–4), 476–495.

McCulloch, Robert E. – Tsay, Ruey S. (1994) Statistical analysis of economic time series via Markov switching models. *Journal of Time Series Analysis*, Vol. 15 (5), 523–539.

McGowan, Carl B. Jr. (2011) An analysis of the technical efficiency of the Russian stock market. *International Business & Economics Research Journal*, Vol. 10 (10), 31–44.

McGowan, Carl B. Jr. – Ibrihim, Izani (2009) An analysis of the day-of-the-week effect in the Russian stock market. *International Business & Economics Research Journal*, Vol. 8 (9), 25–30.

McLean, R. David (2010) Idiosyncratic risk, long-term reversal, and momentum. *Journal of Financial and Quantitative Analysis*, Vol. 45 (4), 883–906.

McLean, R. David – Pontiff, Jeffrey (2016) Does academic research destroy stock return predictability? *The Journal of Finance*, Vol. 71 (1), 5–32.

McNeil, Alexander J. – Frey, Rüdiger (2000) Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach. *Journal of Empirical Finance*, Vol. 7 (3), 271–300.

McNeil, Alexander J. – Frey, Rüdiger – Embrechts, Paul (2005) *Quantitative risk management: Concepts, techniques and tools*. Princeton University Press, Princeton, New Jersey.

Meddahi, Nour (2002) A theoretical comparison between integrated and realized volatility. *Journal of Applied Econometrics*, Vol. 17 (5), 479–508.

Medvedev, Dmitry (2016) Social and economic development of Russia: Finding new dynamics. *Russian Journal of Economics*, Vol. 2 (1), 327–348.

Mehrling, Perry (2000) Minsky and modern finance. *The Journal of Portfolio Management*, Vol. 26 (2), 81–88.

Mei, Xiaoling – Nogales, Francisco J. (2018) Portfolio selection with proportional transaction costs and predictability. *Journal of Banking & Finance*, Vol. 94 (1), 131–151.

Merton, Robert C. (1973) An intertemporal capital asset pricing model. *Econometrica*, Vol. 41 (5), 867–887.

Merton, Robert C. (1974) On the pricing of corporate debt: The risk structure of interest rates. *The Journal of Finance*, Vol. 29 (2), 449–470.

Merton, Robert C. (1980) On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics*, Vol. 8 (1), 323–361.

Merton, Robert C. (1981) On market timing and investment performance I: An equilibrium theory of value for market forecasts. *The Journal of Business*, Vol. 54 (3), 363–406.

Meskarian, Rudabeh – Fliege, Jörg – Xu, Huifu (2014) Stochastic programming with multivariate second order stochastic dominance constraints with application in portfolio optimization. *Applied Mathematics & Optimization*, Vol. 70 (1), 111–140.

Meucci, Attilio (2005) Beyond Black-Litterman: Views on non-normal markets. Working Paper. Social Science Research Network (SSRN). <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=848407>, retrieved 28.9.2017.

Meucci, Attilio (2006) Beyond Black-Litterman in practice: A five-step recipe to input views on non-normal markets. Working Paper. Lehman Brothers, Inc., New York. <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=872577>, retrieved 28.9.2017.

Meucci, Attilio (2007) *Risk and asset allocation*. Corrected 3rd printing. Springer Finance. Springer-Verlag, Berlin; Heidelberg.

Meucci, Attilio (2008a) The Black-Litterman approach: Original model and extensions. Working Paper. Social Science Research Network (SSRN). <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1117574>, retrieved 28.9.2017.

Meucci, Attilio (2008b) Fully flexible views: Theory and practice. Working Paper. Advanced Risk and Portfolio Management (ARPM). <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1213325>, retrieved 28.9.2017.

Meucci, Attilio (2009) Managing diversification. *Risk*, Vol. 22 (5), 74–79.

Meucci, Attilio – Ardia, David – Colasante, Marcello (2014) Portfolio construction and systematic trading with factor entropy pooling. *Risk Magazine*, Vol. 27 (5), 56–61. Available at SSRN. <https://ssrn.com/abstract=1742559>, retrieved 1.5.2020.

Meucci, Attilio – Ardia, David – Keel, Simon (2013) Fully flexible views in multivariate normal markets. Working Paper 13-11. Centre Interuniversitaire sur le Risque, les Politiques Économiques et l'Emploi (CIRPÉE). <https://depot.erudit.org/bitstream/003806dd/1/CIRPEE13-11.pdf>, retrieved 20.1.2021.

Michaud, Richard O. (1989) The Markowitz optimization enigma: Is 'optimized' optimal? *Financial Analysts Journal*, Vol. 45 (1), 31–42.

Michaud, Richard O. (1998) *Efficient asset management: A practical guide to stock portfolio optimization and asset allocation*. Oxford University Press, New York, NY.

Michaud, Richard O. – Esch, David N. – Michaud, Robert O. (2012) Deconstructing Black-Litterman: How to get the portfolio you already knew you wanted. Working Paper. New Frontier. <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2641893>, retrieved 1.8.2021.

Michaud, Richard O. – Michaud, Robert O. (2008) *Efficient asset management: A practical guide to stock portfolio optimization and asset allocation*. Oxford University Press, New York, NY.

Micocci, Marco – Masala, Giovanni (2005) *Backtesting value-at-risk estimation with non-Gaussian marginals*. Working paper. 8th International Congress on Insurance: Mathematics & Economics, Rome.

Mikosch, Thomas – Starica, Catalin (2004) Nonstationarities in financial time series, the long-range dependence, and the IGARCH effects. *The Review of Economics and Statistics*, Vol. 86 (1), 389–390.

Miller, Merton H. (1999) The history of finance. *The Journal of Portfolio Management*, Vol. 25 (4), 95–101.

Modigliani, Franco – Modigliani, Leah (1997) Risk-adjusted performance. *The Journal of Portfolio Management*, Vol. 23 (2), 45–54.

MOEX Russia total return indices. Moscow Exchange. <https://www.moex.com/en/index/totalreturn/MCFTR/lastvalues>, retrieved 17.2.2018.

Moosa, Imad A. (2003) *International financial operations: Arbitrage, hedging, speculation, financing and investment*. Palgrave Macmillan, Basingstoke.

Moritz, Benjamin – Zimmermann, Tom (2016) Tree-based conditional portfolio sorts: The relation between past and future stock returns. Working paper. Ludwig Maximilian University Munich and Harvard University. Social Science Research Network (SSRN). <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2740751>, retrieved 22.5.2020.

Moscow Exchange Sectoral Indices. Moscow Exchange. <http://www.moex.com/en/indices>, retrieved 28.9.2017.

Moskowitz, Tobias J. – Grinblatt, Mark (1999) Do industries explain momentum? *The Journal of Finance*, Vol. 54 (4), 1249–1290.

Moskowitz, Tobias J. – Ooi, Yao Hua – Pedersen, Lasse Heje (2012) Time series momentum. *Journal of Financial Economics*, Vol. 104 (2), 228–250.

Mossin, Jan (1966) Equilibrium in a capital asset market. *Econometrica*, Vol. 34 (4), 768–783.

Muck, Matthias (2010) Trading strategies with partial access to the derivatives market. *Journal of Banking & Finance*, Vol. 34 (6), 1288–1298.

Mullen, Katharine M. – Ardia, David – Gil, David L. – Windover, Donald – Cline, James (2011) DEoptim: An R package for global optimization by differential evolution. *Journal of Statistical Software*, Vol. 40 (6), 1–26.

Musiela, Marek – Rutkowski, Marek (2005) *Martingale methods in financial modelling*. 2nd ed. Volume 36 in Stochastic Modelling and Applied Probability. Springer-Verlag, Berlin.

Nagler, Thomas – Schepsmeier, Ulf – Stoeber, Jakob – Brechmann, Eike Christian – Graeler, Benedikt – Erhardt, Tobias – Almeida, Carlos – Min, Aleksey – Czado, Claudia – Hofmann, Mathias – Killiches, Matthias – Joe, Harry – Vatter, Thibault (2019) Package 'VineCopula': Statistical inference of vine copulas. The Comprehensive R Archive Network (CRAN). <https://cran.r-project.org/web/packages/VineCopula/VineCopula.pdf>, retrieved 1.8.2019.

Nawrocki, David N. (1999) A brief history of downside risk measures. *The Journal of Investing*, Vol. 8 (3), 9–25.

Neftci, Salih N. (2000) Value at risk calculations, extreme events, and tail estimation. *The Journal of Derivatives*, Vol. 7 (3), 23–37.

Negrych, Colin – Senft, Dexter (1989) Portfolio insurance using synthetic puts – the reasons, rewards, and risks. In: *The handbook of fixed income options: Pricing, strategies & applications*, ed. Frank J. Fabozzi, Chapter 12. Probus Publishing, Chicago, IL.

Nelsen, Roger B. (2006) *An introduction to copulas*. 2nd edition. Springer, Portland, OR.

Nelson, Daniel B. (1990) ARCH models as diffusion approximations. *Journal of Econometrics*, Vol. 45 (1), 7–38.

Nelson, Daniel B. (1991) Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, Vol. 59 (2), 347–370.

Nelson, Daniel B. (1992) Filtering and forecasting with misspecified ARCH models I: Getting the right variance with the wrong model. *Journal of Econometrics*, Vol. 52 (2), 61–90.

Nelson, Daniel B. (1994) Asymptotic filtering theory for multivariate ARCH models. Technical Working Paper No. 162. National Bureau of Economic Research, Cambridge, MA. <https://www.nber.org/system/files/working_papers/t0162/t0162.pdf>, retrieved 1.4.2021.

Nelson, Daniel B. – Foster, Dean P. (1994) Asymptotic filtering theory for univariate ARCH models. *Econometrica*, Vol. 62 (1), 1–41.

Nelson, Daniel B. – Foster, Dean P. (1995) Filtering and forecasting with misspecified ARCH models II: Making the right forecast with the wrong model. *Journal of Econometrics*, Vol. 67 (2), 303–335.

Nesterov, Yurii – Nemirovskii, Arkadii (1994) *Interior-point polynomial algorithms in convex programming*. Studies in Applied Mathematics (SIAM), vol. 13. The Society for Industrial and Applied Mathematics, Philadelphia, PA.

Nesvetailova, Anastasia (2016) Against all odds? Russia's financial crisis and global geopolitics. *The International Spectator*, Vol. 51 (1), 113–130.

Newey, Whitney K. – West, Kenneth D. (1987) A simple, positive semidefinite heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, Vol. 56 (1), 203–208.

Nguyen, Duy – Yin, George – Zhang, Qing (2014) A stochastic approximation approach for trend-following trading. In: *Hidden Markov models in finance: Further developments and applications, Volume II*, eds. Rogemar S. Mamon – Robert J. Elliott, 167–184. Springer Science+Business Media, New York, NY.

Nguyen, Tri-Dung – Lo, Andrew W. (2012) Robust ranking and portfolio optimization. *European Journal of Operational Research*, Vol. 221 (2), 407–416.

Nijman, Theo – Swinkels, Laurens – Verbeek, Marno (2004) Do countries or industries explain momentum in Europe? *Journal of Empirical Finance*, Vol. 11 (1), 461–481.

Nivorozhkin, Eugene – Castagneto-Gissey, Giorgio (2016) Russian stock market in the aftermath of the Ukrainian crisis. *Russian Journal of Economics*, Vol. 2 (1), 23–40.

Ono, Shigeki (2017) Financial development and economic growth nexus in Russia. *Russian Journal of Economics*, Vol. 3 (1), 321–332.

Opgen-Rhein, Rainer – Strimmer, Korbinian (2007) Accurate ranking of differentially expressed genes by a distribution-free shrinkage approach. *Statistical Applications in Genetics and Molecular Biology*, Vol. 6 (1), 1–18.

Ortobelli, Sergio – Rachev, Svetlozar T. – Stoyanov, Stoyan – Fabozzi, Frank J. – Biglova, Almira (2005) The proper use of risk measures in portfolio theory. *International Journal of Theoretical and Applied Finance*, Vol. 8 (8), 1107–1133.

Osler, Carol L. (2000) Support for resistance: Technical analysis and intraday exchange rates. *FRBNY (Federal Reserve Bank of New York) Economic Policy Review*, Vol. 6 (2), 53–68.

Osler, Carol L. (2003) Currency orders and exchange rate dynamics: An explanation for the predictive success of technical analysis. *The Journal of Finance*, Vol. 58 (5), 1791–1819.

Pachamanova, Dessislava A. – Fabozzi, Frank J. (2016) *Portfolio construction and analytics*. John Wiley & Sons, Hoboken, NJ.

Pafka, Szilárd – Potters, Marc – Kondor, Imre (2004) Exponential weighting and random-matrix-theory-based filtering of financial covariance matrices for portfolio optimization. Working Paper. arXiv e-prints. <https://arxiv.org/pdf/cond-mat/0402573.pdf>, retrieved 1.8.2021.

Paindaveine, Davy – Bever, Germain, Van (2014) Inference on the shape of elliptical distributions based on the MCD. *Journal of Multivariate Analysis*, Vol. 129 (1), 125–144.

Pak, Olga – Kretzschmar, Gavin Lee (2016) Western sanctions – only half the challenge to Russia's economic union. *Research in International Business and Finance*, Vol. 38 (1), 577–592.

Palomba, Giulio (2008) Multivariate GARCH models and the Black-Litterman approach for tracking error constrained portfolios: An empirical analysis. *Global Business and Economics Review*, Vol. 10 (4), 379–413.

Pastor, Lubos (2000) Portfolio selection and asset pricing models. *The Journal of Finance*, Vol. 55 (1), 179–223.

Pastor, Lubos – Stambaugh, Robert F. (2000) Comparing asset pricing models: An investment perspective. *Journal of Financial Economics*, Vol. 56 (3), 335–381.

Pastor, Lubos – Stambaugh, Robert F. (2003) Liquidity risk and expected stock returns. *The Journal of Political Economy*, Vol. 111 (3), 642–685.

Patton, Andrew J. (2006a) Estimation of multivariate models for time series of possibly different lengths. *Journal of Applied Econometrics*, Vol. 21 (1), 147–173.

Patton, Andrew J. (2006b) Modelling asymmetric exchange rate dependence. *International Economic Review*, Vol. 47 (2), 527–556.

Patton, Andrew (2013) Copula methods for forecasting multivariate time series. In: *Handbook of economic forecasting, Volume 2, Part B*, eds. Graham Elliott – Allan Timmermann, 899–960. North Holland, Amsterdam.

Pedersen, Lasse (2009) When everyone runs for the exit. *International Journal of Central Banking*, Vol. 5 (4), 177–199.

Pedregosa, Fabian – Varoquaux, Gael – Gramfort, Alexandre – Michel, Vincent – Thirion, Bertrand – Grisel, Olivier – Blondel, Mathieu – Prettenhofer, Peter – Weiss, Ron – Dubourg, Vincent – Vanderplas, Jake – Passos, Alexandre – Cournapeau, David – Brucher, Matthieu – Perrot, Matthieu – Duchesnay, Edouard (2011) Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, Vol. 12 (1), 2825–2830.

Pekarek, Edward – Meseha, Maryam (2012) Short selling – the ambrosia or kryptonite of emerging markets? In: *Handbook of short selling*, ed. Greg N. Gregoriou, 365–380. Elsevier, Waltham, MA.

Pentula, S. (1986) Modelling the persistence of conditional variances: A comment. *Econometric Reviews*, Vol. 5 (1), 71–74.

Pérez-Cruz, Fernando – Afonso-Rodríguez, Julio A. – Giner, Javier (2003) Estimating GARCH models using support vector machines. *Quantitative Finance*, Vol. 3 (1), 1–10.

Perold, André F. (1986) *Constant portfolio insurance*. Unpublished manuscript. Harvard Business School, Boston, MA.

Perold, André F. – Sharpe, William F. (1988) Dynamic strategies for asset allocation. *Financial Analyst Journal*, Vol. 51 (1), 16–27.

Peterson, Brian G. – Carl, Peter – Boudt, Kris – Bennett, Ross – Ulrich, Joshua – Zivot, Eric – Lestel, Matthieu – Balkissoon, Kyle – Würtz, Diethelm (2015) Package 'PerformanceAnalytics': Econometric tools for performance and risk analysis. The Comprehensive R Archive Network (CRAN). <https://cran.r-project.org/web/packages/PerformanceAnalytics/PerformanceAnalytics.pdf>, retrieved 25.9.2017.

Peterson, Brian G. – Carl, Peter – Boudt, Kris – Bennett, Ross – Varon, Hezky – Yollin, Guy – Martin, R. Douglas (2018) Package 'PortfolioAnalytics': Portfolio analysis, including numerical methods for optimization of portfolios. The Comprehensive R Archive Network (CRAN). <https://cran.r-project.org/web/packages/PortfolioAnalytics/PortfolioAnalytics.pdf>, retrieved 1.5.2020.

Pézier, Jacques (2007) Global portfolio optimization revisited: A least discrimination alternative to Black-Litterman. ICMA Centre Discussion Papers in Finance. <https://www.icmacentre.ac.uk/discussion-papers/global-portfolio-optimization-revisited-a-least-discrimination-alternative-to-black-litterman>, retrieved 1.8.2021.

Pfaff, Bernhard (2007) Tactical asset allocation: Putting the pieces together. *Slideshow presentation at the 2nd International R/Rmetrics User and Developer Workshop*, Meielisalp, Lake Thune, Switzerland, June 29 – July 3, 2007, 1–54.

Pfaff, Bernhard (2008) *Analysis of integrated and cointegrated time series with R*. 2nd ed. Springer, New York.

Pfaff, Bernhard (2016) *Financial risk modelling and portfolio optimization with R*. 2nd ed. Wiley, Chichester, West Sussex, UK.

Pfaff, Bernhard – Stigler, Matthieu (2018) Package 'vars': VAR modelling. The Comprehensive R Archive Network (CRAN). <https://cran.r-project.org/web/packages/vars/vars.pdf>, retrieved 1.7.2020.

Pfaff, Bernhard – Vandenberghe, Lieven – Andersen, Martin – Dahl, Joachim (2020) Package 'cccp': Cone constrained convex problems. The Comprehensive R Archive Network (CRAN). <https://cran.r-project.org/web/packages/cccp/cccp.pdf>, retrieved 16.2.2021.

Pfaff, Bernhard – Zivot, Eric – McNeil, Alexander – Stephenson, Alec (2018) Package 'evir': Extreme values in R. The Comprehensive R Archive Network (CRAN). <https://cran.r-project.org/web/packages/evir/evir.pdf>, retrieved 1.7.2020.

Pfaff, Bernhard – Zivot, Eric – Stigler, Matthieu (2016) Package 'urca': Unit root and cointegration tests for time series data. The Comprehensive R Archive Network (CRAN).<https://cran.r-project.org/web/packages/urca/urca.pdf>, retrieved 1.7.2020.

Phelan, Michael J. (1997) Probability and statistics applied to the practice of financial risk management: The case of J.P. Morgan's RiskMetrics™. *Journal of Financial Services Research*, Vol. 12 (2/3), 175–200.

Philips, Thomas K. – Rogers, Greg T. – Capaldi, Robert E. (1996) Tactical asset allocation: 1977-1994. *The Journal of Portfolio Management*, Vol. 23 (1), 57–64.

Phillips, Don – Lee, Joan (1989) Differentiating tactical asset allocation from market timing. *Financial Analysts Journal*, Vol. 45 (2), 14–16.

Pliska, Stanley R. – Suzuki, Kiyoshi (2004) Optimal tracking for asset allocation with fixed and proportional transaction costs. *Quantitative Finance*, Vol. 4 (2), 233–243.

Pojarliev, Momtchil – Polasek, Wolfgang (2003) Portfolio construction by volatility forecasts: Does the covariance structure matter? *Financial Markets and Portfolio Management*, Vol. 17 (1), 103–116.

Poon, Ser-Huang (2005) *A practical guide to forecasting financial market volatility*. John Wiley & Sons Ltd, Chichester, England.

Poon, Ser-Huang – Granger, Clive W. J. (2003) Forecasting volatility in financial markets: A review. *Journal of Economic Literature*, Vol. 41 (2), 478–539.

Portfolio optimization theory. Financial Toolbox™ Documentation. The MathWorks®, Inc. <https://se.mathworks.com/help/finance/portfolio-optimization-theory-mad.html#btyn1fe-7> , retrieved 28.5.2021.

Prigent, Jean-Luc (2007) *Portfolio optimization and performance analysis*. Chapman & Hall/CRC, Boca Raton, FL.

Pätäri, Eero – Luukka, Pasi – Fedorova, Elena – Garanina, Tatiana (2017) The anatomy of returns from moving average trading rules in the Russian stock market. *Applied Economic Letters*, Vol. 24 (5), 311–318.

Qian, Edward – Gorman, Stephen (2001) Conditional distribution in portfolio theory. *Financial Analysts Journal*, Vol. 57 (2), 44–51.

Qian, Edward (2003) Tactical asset allocation with pairwise strategies: Using pairwise information to influence weights. *The Journal of Portfolio Management*, Vol. 30 (1), 39–48.

Qian, Edward (2005) Risk parity portfolios: Efficient portfolios through true diversification. White paper. PanAgora, Boston, MA. <https://www.panagora.com/assets/PanAgora-Risk-Parity-Portfolios-Efficient-Portfolios-Through-True-Diversification.pdf>, retrieved 20.7.2019.

R Core Team (2018) R: A language and environment for statistical computing. <https://www.r-project.org/>, retrieved 17.2.2018.

Rab, Nikolaus – Warnung, Richard (2011) Scaling portfolio volatility and calculating risk contributions in the presence of serial cross-correlations. Working Paper. arXiv Preprint. <https://arxiv.org/pdf/1009.3638.pdf>, retrieved 1.8.2021.

Rachev, Svetlozar – Jašić, Teo – Stoyanov, Stoyan – Fabozzi, Frank J. (2007) Momentum strategies based on reward–risk stock selection criteria. *Journal of Banking & Finance*, Vol. 31 (8), 2325–2346.

Rachev, Svetlozar – Ortobelli, Sergio – Stoyanov, Stoyan – Fabozzi, Frank J. – Biglova, Almira (2008) Desirable properties of an ideal risk measure in portfolio theory. *International Journal of Theoretical and Applied Finance*, Vol. 11 (1), 19–54.

Raffinot, Thomas (2018) Hierarchical clustering-based asset allocation. *The Journal of Portfolio Management, Multi-Asset Special Issue*, Vol. 44 (2), 89–99.

Ranaldo, Angelo – Favre, Laurent (2005) How to price hedge funds: From two- to four-moment CAPM. Working Paper. <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=474561>, retrieved 1.4.2021.

Ratinov, Mikhail (1997) Investing in Russian securities: Analysis of capital market development. *Fordham International Law Journal*, Vol. 21 (5), 1772–1801.

Rebonato, Riccardo – Denev, Alexander (2013) *Portfolio management under stress: A Bayesian-net approach to coherent asset allocation*. Cambridge University Press, Cambridge, United Kingdom.

Rey, David (2004a) Tactical asset allocation: An alternative definition. Working Paper No. 3/04. University of Basel (WWZ), Department of Finance, Basel. <https://www.econbiz.de/archiv1/2008/59446_tactical_asset_allocation.pdf> , retrieved 1.8.2021.

Rey, David (2004b) Tactical asset allocation, parameter uncertainty, and intertemporal hedging: An exploratory study for the Swiss stock market. Working Paper No. 4/04. University of Basel (WWZ), Department of Finance, Basel. <https://www.econbiz.de/archiv1/2008/59445_tactical_asset_allocation.pdf>, retrieved 1.8.2021.

Riedel, Frank (2004) Dynamic coherent risk measures. *Stochastic Processes and their Applications*, Vol. 112 (2), 185–200.

Ripley, Brian – Venables, Bill – Bates, Douglas M. – Hornik, Kurt – Gebhardt, Albrecht – Firth, David (2021) Package 'MASS': Support functions and datasets for Venables and Ripley's MASS. The Comprehensive R Archive Network (CRAN). <https://cran.r-project.org/web/packages/MASS/MASS.pdf>, retrieved 25.2.2021.

Roberge, Mathieu – Le Moigne, Cécile (2005) A multivariate dichotomic approach for tactical asset allocation. *Journal of Asset Management*, Vol. 6 (3), 206–218.

Rockafellar, R. Tyrrell – Uryasev, Stanislav (2002) Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance*, Vol. 26 (1), 1443–1471.

Rocke, David M. (1996) Robustness properties of S-estimators of multivariate location and shape in high dimension. *The Annals of Statistics*, Vol. 24 (3), 1327–1345.

Rockinger, Michael – Urga, Giovanni (2000) The evolution of stock markets in transition economies. *Journal of Comparative Economics*, Vol. 28 (1), 456–472.

Rockinger, Michael – Jondeau, Eric (2001) *Conditional dependency of financial series: An application of copulas*. Notes d'Études et de Recherche NER 82, Banque de France, Paris.

Roncalli, Thierry (2010) Understanding the impact of weights constraints in portfolio theory. Technical Paper. Lyxor Asset Management, Paris. <http://www.thierry-roncalli.com/download/shrinkage.pdf>, retrieved 25.2.2021.

Roncalli, Thierry (2013) Introducing expected returns into risk parity portfolios: A new framework for asset allocation. Working Paper. Social Science Research Network (SSRN). <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2321309>, retrieved 1.11.2021.

Roncalli, Thierry (2014) *Introduction to risk parity and budgeting*. Chapman & Hall/CRC Financial Mathematics Series. CRC Press, Boca Raton, FL.

Ross, Stephen A. (1976) The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, Vol. 13 (3), 341–360.

Rousseeuw, Peter J. (1984) Least median of squares regression. *Journal of the American Statistical Association*, Vol. 79 (388), 871–880.

Rousseeuw, Peter (1985) Multivariate estimation with high breakdown point. In: *Mathematical statistics and applications: Volume B*, eds. W. Grossmann – G. Pflug – I. Vincze – W. Wertz, 283–297. Reidel Publishing Company, Dordrecht.

Rousseeuw, Peter J. – Leroy, Annick M. (1987) *Robust regression and outlier detection*. John Wiley & Sons, New York, NY.

Rousseeuw, Peter J. – Van Driessen, Katrien (1999) A fast algorithm for the minimum covariance determinant estimator. *Technometrics*, Vol. 41 (3), 212–223.

Rouwenhorst, K. Geert (1998a) International momentum strategies. *The Journal of Finance*, Vol. 53 (1), 267–284.

Rouwenhorst, K. Geert (1998b) Local return factors and turnover in emerging stock markets. Discussion paper. <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=115788>, retrieved 28.10.2018.

Rouwenhorst, K. Geert (1999) Local return factors and turnover in emerging stock markets. *The Journal of Finance*, Vol. 54 (4), 1439–1464.

Roy, Andrew D. (1952) Safety-first and the holding of assets. *Econometrica*, Vol. 20 (1), 431–449.

Rozhkov, A. (2005) Forecasting Russian stock market trends. *Problems of Economic Transition*, Vol. 48 (6), 48–65.

Rubinstein, Mark – Leland, Hayne E. (1981) Replicating options with positions in stock and cash. *Financial Analysts Journal*, Vol. 37 (4), 63–72.

Ruppert, David – Matteson, David S. (2015) *Statistics and data analysis for financial engineering: With R examples*. 2nd ed. Springer-Verlag, New York, NY.

Sadka, Ronnie (2006) Momentum and post-earnings-announcement drift anomalies: The role of liquidity risk. *Journal of Financial Economics*, Vol. 80 (2), 309–349.

Sahamkhadam, Maziar – Stephan, Andreas – Östermark, Ralf (2018) Portfolio optimization based on GARCH-EVT-Copula forecasting models. *International Journal of Forecasting*, Vol. 34 (3), 497–506.

Saleem, Kashif (2009) International linkage of the Russian market and the Russian financial crisis: A multivariate GARCH analysis. *Research in International Business and Finance*, Vol. 23 (1), 243–256.

Saleem, Kashif (2011) Time varying correlations between stock and bond returns: Empirical evidence from Russia. *Asian Journal of Finance & Accounting*, Vol. 3 (1), 72–85.

Saleem, Kashif (2014) Modeling long memory in the Russian stock market: Evidence from major sectoral indices. *The Journal of Applied Business Research*, Vol. 30 (2), 567–574.

Saleem, Kashif – Vaihekoski, Mika (2008) Pricing of global and local sources of risk in Russian stock market. *Emerging Markets Review*, Vol. 9 (1), 40–56.

Saleem, Kashif – Vaihekoski, Mika (2010) Time-varying global and local sources of market and currency risks in Russian stock market. *International Review of Economics and Finance*, Vol. 19 (1), 686–697.

Salibián-Barrera, Matías – Van Aelst, Stefan – Willems, Gert (2006) Principal components analysis based on multivariate MM estimators with fast and robust bootstrap. *Journal of the American Statistical Association*, Vol. 101 (475), 1198–1211.

Salibian-Barrera, Matías – Yohai, Víctor J. (2006) A fast algorithm for S-regression estimates. *Journal of Computational and Graphical Statistics*, Vol. 15 (2), 414–427.

Santiago, Rafael, De – Estrada, Javier (2013) Geometric mean maximization: Expected, observed, and simulated performance. *The Journal of Investing*, Vol. 22 (2), 109–119.

Santis, Giorgio, De – Litterman, Bob – Vesval, Adrien – Winkelmann, Kurt (2003) Covariance matrix estimation. In: *Modern investment management: An equilibrium approach*, ed. Bob Litterman, 224–248. John Wiley & Sons, Hoboken, NJ.

Santos, Alejandro (2003) Debt crisis in Russia: The road from default to sustainability. In: *Russia rebounds*, eds. David Owen – David O. Robinson, 154–183. International Monetary Fund, Washington, DC.

Satchell, Stephen – Scowcroft, Alan (2000) A demystification of the Black-Litterman model: Managing quantitative and traditional portfolio construction. *Journal of Asset Management*, Vol. 1 (2), 138–150.

Scherer, Bernd (2002) Portfolio resampling: Review and critique. *Financial Analysts Journal*, Vol. 58 (6), 98–109.

Scherer, Bernd (2007) Can robust portfolio optimisation help to build better portfolios? *Journal of Asset Management*, Vol. 7 (1), 374–387.

Scherer, Bernd (2010) *Portfolio construction and risk budgeting*. 4th ed. Risk Books, London.

Scherer, Bernd – Martin, Douglas (2005) *Introduction to modern portfolio optimization with NuOPT, S-PLUS and S+Bayes*. Springer Science+Business Media, New York, NY.

Schmidbauer, Harald – Rösch, Angi – Uluceviz, Erhan – Erkol, Narod (2016) The Russian stock market during the Ukrainian crisis: A network perspective. *Czech Journal of Economics and Finance (Finance a uver)*, Vol. 66 (6), 478–509.

Schmittmann, Jochen M. (2010) Currency hedging for international portfolios. IMF Working Paper WP/10/151, International Monetary Fund. <https://www.researchgate.net/publication/46433354_Currency_Hedging_for_International_Portfolios>, retrieved 1.8.2021.

Schwert, G. William (1990) Stock volatility and the crash of '87. *Review of Financial Studies*, Vol. 3 (1), 77–102.

Schwert, G. William (2003) Anomalies and market efficiency. In: *Handbook of the economics of finance. Volume 1, Part B*, eds. George M. Constantinides – Milton Harris – René M. Stulz, 939–974. North Holland, Amsterdam.

Schäfer, Juliane – Strimmer, Korbinian (2005) A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. *Statistical Applications in Genetics and Molecular Biology*, Vol. 4 (1), 1–32.

Schäfer, Juliane – Opgen-Rhein, Rainer – Strimmer, Korbinian – Zuber, Verena – Ahdesmäki, Miika – Silva, A. Pedro Duarte (2017) Package 'corpcor': Efficient estimation of covariance and (partial) correlation. The Comprehensive R Archive Network (CRAN). <https://cran.r-project.org/web/packages/corpcor/corpcor.pdf>, retrieved 1.4.2018.

Schöttle, Katrin (2007) *Robust optimization with application in asset management*. Doctoral dissertation. Technische Universität München, München.

Schöttle, Katrin – Werner, Ralf (2009) Robustness properties of mean-variance portfolios. *Optimization*, Vol. 58 (6), 641–663.

Scott, James – Stumpp, Margaret – Xu, Peter (2003) News, not trading volume, builds momentum. *Financial Analysts Journal*, Vol. 59 (2), 45–54.

Scott, Robert C. – Horvath, Philip A. (1980) On the direction of preference for moments of higher order than the variance. *The Journal of Finance*, Vol. 35 (4), 915–919.

Sehgal, Ruchika – Mehra, Aparna (2021) Robust reward–risk ratio portfolio optimization. *International Transactions in Operational Research*, Vol. 28 (1), 2169–2190.

Shankar, S. Gowri – Miller, James M. (2006) Market reaction to changes in the S&P SmallCap 600 index. *The Financial Review*, Vol. 41 (3), 339–360.

Shapiro, Alexander (2013) On Kusuoka representation of law invariant risk measures. *Mathematics of Operations Research*, Vol. 38 (1), 142–152.

Sharpe, Willian F. (1963) A simplified model for portfolio analysis. *Management Science*, Vol. 9 (1), 603–621.

Sharpe, William F. (1964) Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, Vol. 19 (1), 425–442.

Sharpe, William F. (1994) The Sharpe ratio. *The Journal of Portfolio Management*, Vol. 21 (1), 49–58.

Sharpe, William F. – Alexander, Gordon J. – Bailey, Jeffery V. (1999) *Investments*. 6^{th} ed. Prentice Hall, Upper Saddle River, NJ.

Shaw, William T. (2011) Portfolio optimization for VaR, CVaR, Omega and utility with general return distributions: A Monte Carlo approach for long-only and bounded short portfolios with optional robustness and a simplified approach to covariance matching. Working Paper. University College London. <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1856476>, retrieved 16.2.2021.

Shefrin, Hersh – Statman, Meir (1985) The disposition to sell winners too early and ride losers too long: Theory and evidence. *The Journal of Finance*, Vol. 40 (3), 777–790.

Shum, Pauline – Hejazi, Walid – Haryanto, Edgar – Rodier, Arthur (2015) Intraday share price volatility and leveraged ETF rebalancing. *Review of Finance*, Vol. 20 (6), 2379–2409.

Siami-Namini, Sima – Siami-Namin, Akbar (2018) Forecasting economic and financial time series: ARIMA vs. LSTM. Working Paper. Department of Agricultural and Applied Economics, Texas Tech University, Lubbock, TX; Department of Computer Science, Texas Tech University, Lubbock, TX. ArXiv. <https://arxiv.org/ftp/arxiv/papers/1803/1803.06386.pdf>, retrieved 15.6.2020.

Silva, Alexandre S., da – Lee, Wai – Pornrojnangkool, Bobby (2009) The Black–Litterman model for active portfolio management. *The Journal of Portfolio Management*, Vol. 35 (1), 61–70.

Simola, Heli – Solanko, Laura (2017) Overview of Russia's oil and gas sector. BOFIT Policy Brief 5/2017. Bank of Finland, Institute for Economies in Transition. <https://helda.helsinki.fi/bof/bitstream/handle/123456789/14701/bpb0517.pdf>, retrieved 15.8.2021.

Simonian, Joseph (2014) Copula-opinion pooling with complex opinions. *Quantitative Finance*, Vol. 14 (6), 941–946.

Sinclair, Euan (2010) *Option trading: Pricing and volatility strategies and techniques*. John Wiley & Sons, Hoboken, NJ.

Sinclair, Euan (2013) *Volatility trading*. 2^{nd} ed. John Wiley & Sons, Hoboken, NJ.

Singh, Abhay K. – Allen, David E. – Powell, Robert J. (2013) Extreme market risk and extreme value theory. *Mathematics and Computers in Simulation*, Vol. 94 (1), 310–328.

Singh, Abhay K. – Allen, David E. – Powell, Robert J. – Reddy, Krishna (2014) Multivariate financial dependence analysis of Asian markets using vine copulas. Social Science Research Network (SSRN). <https://ssrn.com/abstract=2449932>, retrieved 1.8.2019.

Singh, Abhay Kumar – Allen, David Edmund (2017) *R in finance and economics: A beginner's guide*. World Scientific Publishing Company, Singapore.

Singh, Manoj K. (1997) Value at risk using principal components analysis. *The Journal of Portfolio Management*, Vol. 24 (1), 101–112.

Smirnov, Sergey (2015) Economic fluctuations in Russia (from the late 1920s to 2015). *Russian Journal of Economics*, Vol. 1 (1), 130–153.

Song, Yongsheng – Yan, Jia-An (2009) Risk measures with comonotonic subadditivity or convexity and respecting stochastic orders. *Insurance: Mathematics and Economics*, Vol. 45 (3), 459–465.

Sortino, Frank A. – Price, Lee N. (1994) Performance measurement in a downside risk framework. *The Journal of Investing*, Vol. 3 (3), 59–64.

Stahel, Werner A. (1987) Estimation of a covariance matrix with location: Asymptotic formulas and optimal B-robust estimators. *Journal of Multivariate Analysis*, Vol. 22 (2), 296–312.

Stambaugh, Robert F. (1997) Analyzing investments whose histories differ in length. *Journal of Financial Economics*, Vol. 45 (3), 285–331.

Starica, Catalin (2003) Is GARCH(1,1) as good a model as the Nobel prize accolades would imply? Working Paper. Social Science Research Network (SSRN). <https://ssrn.com/abstract=637322>, retrieved 1.8.2019.

Stein, Michael – Füss, Roland – Drobetz, Wolfgang (2009) Fixed-income portfolio allocation including hedge fund strategies: A copula opinion pooling approach. *The Journal of Fixed Income*, Vol. 18 (4), 78–91.

Stigler, George J. (1965) *Essays in the history of economics*. University of Chicago Press, Chicago, IL.

Stoyanov, S. V. – Rachev, S. T. – Fabozzi, F. J. (2007) Optimal financial portfolios. *Applied Mathematical Finance*, Vol. 14 (5), 401–436.

Sun, Yiguo – Wu, Ximing (2018) Leverage and volatility feedback effects and conditional dependence index: A nonparametric study. *Journal of Risk and Financial Management*, Vol. 11 (2), 1–20.

Szegö, Giorgio (ed.) (2004) *Risk measures for the 21st century*. John Wiley & Sons Ltd, Somerset, NJ.

Tabata, Shinichiro (2009) The impact of global financial crisis on the mechanism of economic growth in Russia. *Eurasian Geography and Economics*, Vol. 50 (6), 682–698.

Takeuchi, Lawrence – Lee, Yu-Ying (2013) Applying deep learning to enhance momentum trading strategies in stocks. Working paper. Stanford University. <http://cs229.stanford.edu/proj2013/TakeuchiLee-Applying-DeepLearningToEnhanceMomentumTradingStrategiesInStocks.pdf>, retrieved 20.5.2020.

Taleb, Nassim Nicholas (2010) *The black swan: The impact of the highly improbable*. 2nd ed. Random House Trade Paperbacks, New York, NY.

Tatsuoka, Kay S. – Tyler, David E. (2000) On the uniqueness of S-functionals and M-functionals under nonelliptical distributions. *The Annals of Statistics*, Vol. 28 (4), 1219–1243.

Tay, Francis E. H. – Cao, Lijuan (2001) Application of support vector machines in financial time series forecasting. *Omega*, Vol. 29 (4), 309–317.

Taylor, Stephen J. (1986) *Modelling financial time series*. John Wiley, Chichester, England.

Tee, Chyng Wen – Huang, Shirley – Lim, Kian Guan (2017) Performance control and risk calibration in the Black–Litterman model. *The Journal of Portfolio Management*, Vol. 43 (3), 126–135.

Teplova, Tamara – Shutova, Evgeniya (2011) A higher moment downside framework for conditional and unconditional CAPM in the Russian stock market. *Eurasian Economic Review*, Vol. 1 (2), 157–178.

Thaler, Richard H. (1999) Mental accounting matters. *Journal of Behavioral Decision Making*, Vol. 12 (3), 183–206.

Theussl, Stefan – Hornik, Kurt – Buchta, Christian – Schwendinger, Florian – Schuchardt, Heinrich (2019) Package 'Rglpk': R/GNU linear programming kit interface. The Comprehensive R Archive Network (CRAN). <https://cran.r-project.org/web/packages/Rglpk/Rglpk.pdf>, retrieved 3.9.2020.

Thomson Reuters Global Equity Indices. Index Methodology. October 2015. Thomson Reuters. <https://www.thomsonreuters.com/content/dam/openweb/documents/pdf/tr-com-financial/methodology/global-equity-index-methodology-oct-2015.pdf>, retrieved 9.8.2019.

Timmermann, Allan – Granger, Clive W. J. (2004) Efficient market hypothesis and forecasting. *International Journal of Forecasting*, Vol. 20 (1), 15–27.

Todorov, Valentin (2020) Package 'rrcov': Scalable robust estimators with high breakdown point. The Comprehensive R Archive Network (CRAN). <https://cran.r-project.org/web/packages/rrcov/rrcov.pdf>, retrieved 16.2.2021.

Todorov, Valentin – Filzmoser, Peter (2009) An object-oriented framework for robust multivariate analysis. *Journal of Statistical Software*, Vol. 32 (3), 1–47.

Tokat, Yesim – Wicas, Nelson (2007) Portfolio rebalancing in theory and practice. *The Journal of Investing*, Vol. 16 (2), 52–59.

Tokat, Yesim – Wicas, Nelson – Stockton, Kimberly (2007) Practical guidelines for tactical asset allocation strategy evaluation. *The Journal of Investing*, Vol. 16 (3), 33–41.

Topaloglou, Nikolas – Vladimirou, Hercules – Zenios, Stavros A. (2011) Optimizing international portfolios with options and forwards. *Journal of Banking & Finance*, Vol. 35 (1), 3188–3201.

Total Return Indices. Moscow Exchange. <http://www.moex.com/en/index/totalreturn.aspx>, retrieved 13.10.2017.

Trapletti, Adrian – Hornik, Kurt – LeBaron, Blake (2020) Package 'tseries': Time series analysis and computational finance. The Comprehensive R Archive Network (CRAN). <https://cran.r-project.org/web/packages/tseries/tseries.pdf>, retrieved 8.5.2021.

Tsai, Henghshiu – Chan, K. S. (2005) Maximum likelihood estimation of linear continuous time long memory processes with discrete time data. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, Vol. 67 (5), 703–716.

Tsantekidis, Avraam – Passalis, Nikolaos – Tefas, Anastasios – Kanniainen, Juho – Gabbouj, Moncef – Iosifidis, Alexandros (2017) Forecasting stock prices from the limit order book using convolutional neural networks. *Paper presented at the 19th IEEE International Conference on Business Informatics (CBI)*, Thessaloniki, Greece, July 24–27, 2017, 7–12.

Tsay, Ruey S. (2010) *Analysis of financial time series*. 3rd ed. Wiley, Hoboken, NJ.

Tukey, John W. (1962) The future of data analysis. *The Annals of Mathematical Statistics*, Vol. 33 (1), 1–67.

Tütüncü, R. H. – Koenig, M. (2004) Robust asset allocation. *Annals of Operations Research*, Vol. 132 (1), 157–187.

Tütüncü, R. H. – Toh, K. C. – Todd, M. J. (2003) Solving semidefinite-quadratic-linear programs using SDPT3. *Mathematical Programming*, Vol. 95 (2), 189–217.

Uryasev, Stan (2013) Case study: Portfolio optimization with probabilistic constraint and fixed and proportional transaction costs. Stony Brook University, New York. <http://uryasev.ams.stonybrook.edu/index.php/research/testproblems/financial_engineering/portfolio-optimization-with-probabilistic-constraint-and-fixed-and-proportional-transaction-costs/>, retrieved 12.3.2021.

Value-at-Risk estimation and backtesting. MATLAB® & Simulink® Example. MathWorks Nordic. <https://se.mathworks.com/help/risk/value-at-risk-estimation-and-backtesting.html>, retrieved 1.4.2021.

Van der Hart, Jaap – Slagter, Erica – van Dijk, Dick (2003) Stock selection strategies in emerging markets. *Journal of Empirical Finance*, Vol. 10 (1), 105–132.

Vander Weide, James H. (2010) Principles for lifetime portfolio selection: Lessons from portfolio theory. In: *Handbook of portfolio construction: Contemporary applications of Markowitz techniques*, ed. John B. Guerard, Jr., 153–177. Springer Science+Business Media, New York, NY.

Vanteeva, Nadia (2016) In the absence of private property rights: Political control and state corporatism during Putin's first tenure. *Russian Journal of Economics*, Vol. 2 (1), 41–55.

Vayanos, Dimitri – Woolley, Paul (2013) An institutional theory of momentum and reversal. Working paper. London School of Economics. <https://personal.lse.ac.uk/vayanos/Papers/ITMR_RFS13.pdf>, retrieved 23.5.2019.

Veebel, Viljar – Markus, Raul (2016) At the dawn of a new era of sanctions: Russian-Ukrainian crisis and sanctions. *Orbis*, Vol. 60 (1), 128–139.

Velicer, Wayne F. (1976) Determining the number of components from the matrix of partial correlations. *Psychometrika*, Vol. 41 (3), 321–327.

Venkataraman, Subu (1997) Value at risk for a mixture of normal distributions: The use of quasi-Bayesian estimation techniques. *Economic Perspectives (Federal Reserve Bank of Chicago)*, Vol. 21 (2), 2–13.

Vesterstrom, Jakob – Thomsen, René (2004) A comparative study of differential evolution, particle swarm optimization, and evolutionary algorithms on numerical benchmark problems. In: *Proceedings of the 2004 Congress on Evolutionary Computation*, Portland, OR, USA, June 19–23, 2004, 1980–1987.

Vizgunov, Arseniy – Goldengorin, Boris – Kalyagin, V. – Koldanov, P. – Koldanov, Petr A. – Pardalos, Panos M. (2014) Network approach for the Russian stock market. *Computational Management Science*, Vol. 11 (1-2), 45–55.

Voskoboynikov, Ilya B. (2017) Sources of long run economic growth in Russia before and after the global financial crisis. *Russian Journal of Economics*, Vol. 3 (1), 348–365.

Walters, Jay (2014) The Black-Litterman model in detail. Social Science Research Network (SSRN). <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1314585>, retrieved 23.5.2019.

Wang, Peng – Kochard, Larry (2012) Using a Z-score approach to combine value and momentum in tactical asset allocation. *The Journal of Wealth Management*, Vol. 15 (1), 52–71.

Wang, Wei (2010) Strategic asset allocation and Markov regime switch with GARCH model. *Proceedings of the Northeast Business & Economics Association*, Vol. 1 (1), 391–399.

Wang, Zong-Run – Chen, Xiao-Hong – Jin, Yan-Bo – Zhou, Yan-Ju (2010) Estimating risk of foreign exchange portfolio: Using VaR and CVaR based on GARCH-EVT-copula model. *Physica A: Statistical Mechanics and its Applications*, Vol. 389 (21), 4918–4928.

Wang, Zongrun – Jin, Yanbo – Zhou, Yanju (2010) Estimating portfolio risk using GARCH-EVT-copula model: An empirical study on exchange rate market. In: *Advances in neural network research and applications*, eds. Zhigang Zeng – Jun Wang, 65–72. Springer, Berlin.

Weber, Stefan (2003) Distribution-invariant dynamic risk measures. Working Paper. Humboldt-Universität zu Berlin, Institut für Mathematic, Berlin. <https://core.ac.uk/download/pdf/127602179.pdf>, retrieved 1.4.2021.

Weber, Stefan (2006) Distribution-invariant risk measures, information, and dynamic consistency. *Mathematical Finance*, Vol. 16 (2), 419–441.

Weigel, Eric J. (1991) The performance of tactical asset allocation. *Financial Analysts Journal*, Vol. 47 (5), 63–70.

Welch, B. L. (1947) The generalization of 'Student's' problem when several different population variances are involved. *Biometrika*, Vol. 34 (1), 28–35.

Welsch, Roy E. – Zhou, Xinfeng (2007) Application of robust statistics to asset allocation models. *REVSTAT – Statistical Journal*, Vol. 5 (1), 97–114.

Weron, Rafal (2002) Estimating long-range dependence: Finite sample properties and confidence intervals. *Physica A: Statistical Mechanics and Its Applications*, Vol. 312 (2), 285–299.

Wilcox, Rand (2012) *Introduction to robust estimation and hypothesis testing*. 3rd ed. Academic Press, Waltham, MA.

Wilmott, Paul (1998) *Derivatives: The theory and practice of financial engineering*. John Wiley & Sons, Chichester, West Sussex.

Wilson, Andrew Gordon – Ghahramani, Zoubin (2010) Copula processes. In: *Advances in Neural Information Processing Systems 23 (NIPS 2010)*, eds. J. D. Lafferty – C. K. I. Williams – J. Shawe-Taylor – R. S. Zemel – A. Culotta, Vancouver, Canada, December 6–11, 2010, 1–9.

Würtz, Diethelm – Chalabi, Yohan – Chen, William – Ellis, Andrew (2009) *Portfolio optimization with R/Rmetrics*. Rmetrics Association & Finance Online GmbH, Zurich, Switzerland.

Würtz, Diethelm – Setz, Tobias – Chalabi, Yohan (2017a) Package 'fArma': Rmetrics – modelling ARMA time series processes. The Comprehensive R Archive Network (CRAN). <http://cran.nexr.com/web/packages/fArma/fArma.pdf>, retrieved 8.5.2021.

Würtz, Diethelm – Setz, Tobias – Chalabi, Yohan (2017b) Package 'fExtremes': Rmetrics – modelling extreme events in finance. The Comprehensive R Archive Network (CRAN). <https://cran.r-project.org/web/packages/fExtremes/fExtremes.pdf>, retrieved 1.8.2019.

Würtz, Diethelm – Setz, Tobias – Chalabi, Yohan (2020) Package 'fMultivar': Rmetrics – analysing and modeling multivariate financial return distributions. The Comprehensive R Archive Network (CRAN). <https://cran.r-project.org/web/packages/fMultivar/fMultivar.pdf>, retrieved 1.7.2020.

Würtz, Diethelm – Setz, Tobias – Chalabi, Yohan – Boudt, Chris – Chausse, Pierre – Miklovac, Michal (2020) Package 'fGarch': Rmetrics – autoregressive conditional heteroskedastic modelling. The Comprehensive R Archive Network (CRAN). <https://cran.r-project.org/web/packages/fGarch/fGarch.pdf>, retrieved 1.7.2020.

Würtz, Diethelm – Setz, Tobias – Chalabi, Yohan – Chen, William – Ellis, Andrew (2015a) *Portfolio optimization with R/Rmetrics*. Rmetrics Association and Finance Online GmbH, Zurich, Switzerland.

Würtz, Diethelm – Setz, Tobias – Chalabi, Yohan – Rmetrics Core Team (2015b) Package 'fPortfolio': Rmetrics – portfolio selection and optimization. The Comprehensive R Archive Network (CRAN). <https://cran.r-project.org/web/packages/fPortfolio/fPortfolio.pdf>, retrieved 1.6.2017.

Xue, Jingming – Zhou, SiHang – Liu, Qiang – Liu, Xinwang – Yin, Jianping (2018) Financial time series prediction using $\ell_{2,1}$RF-ELM. *Neurocomputing*, Vol. 277 (14), 176–186.

Yamai, Yasuhiro – Yoshiba, Toshinao (2002) Comparative analyses of expected shortfall and value-at-risk: Their estimation error, decomposition, and optimization. *Monetary and Economic Studies*, Vol. 20 (1), 87–121.

Yang, Jen-Wei – Tsai, Shu-Yu – Shyu, So-De – Chang, Chia-Chien (2016) Pairs trading: The performance of a stochastic spread model with regime switching–evidence from the S&P 500. *International Review of Economics and Finance*, Vol. 43 (1), 139–150.

Yang, Junmin – Cao, Zhiguang – Han, Qiheng – Wang, Qiyu (2018) Tactical asset allocation on technical trading rules and data snooping. *Pacific-Basin Finance Journal*, Vol. 57 (1), 1–15.

Yang, Xuebing – Zhang, Huilan (2019) Extreme absolute strength of stocks and performance of momentum strategies. *Journal of Financial Markets*, Vol. 44 (1), 71–90.

Ye, Kai – Parpas, Panos – Rustem, Berç (2012) Robust portfolio optimization: A conic programming approach. *Computational Optimization and Applications*, Vol. 52 (2), 463–481.

Yeh, Chi-Yuan – Huang, Chi-Wei – Lee, Shie-Jue (2011) A multiple-kernel support vector regression approach for stock market price forecasting. *Expert Systems with Applications*, Vol. 38 (3), 2177–2186.

Yin, Chenyang – Perchet, Romain – Soupé, François (2019) A practical guide to robust portfolio optimization. Working Paper. BNB Paribas Asset Management. <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3490680>, retrieved 4.3.2020.

Yohai, Viktor J. (1987) High breakdown-point and high efficiency robust estimates for regression. *The Annals of Statistics*, Vol. 15 (2), 642–656.

Yohai, Viktor J. – Stahel, Werner A. – Zamar, Ruben H. (1991) A procedure for robust estimation and inference in linear regression. In: *Directions in robust statistics and diagnostics: Part II (The IMA volumes in mathematics and its applications, vol. 34)*, eds. Werner Stahel – Sanford Weisberg, 365–374. Springer-Verlag, New York, NY.

Zakamouline, Valeri I. (2002) *Optimal portfolio selection with transaction costs for a CARA investor with finite horizon*. Research paper. Institute of Finance and Management Science, Norwegian School of Economics and Business Administration, Bergen.

Zakoian, Jean-Michel (1994) Threshold heteroskedastic models. *Journal of Economic Dynamics and Control*, Vol. 18 (5), 931–955.

Zhang, Bangzheng – Wei, Yu – Yu, Jiang – Lai, Xiaodong – Peng, Zhenfeng (2014) Forecasting VaR and ES of stock index portfolio: A Vine copula method. *Physica A: Statistical Mechanics and its Applications*, Vol. 416 (3), 112–124.

Zhang, Dan – Wei, Bin (2014) Comparison between differential evolution and particle swarm optimization algorithms. In: *Proceedings of 2014 IEEE International Conference on Mechatronics and Automation (ICMA)*, Tianjin, China, August 3–6, 2014, 239–244.

Zhang, G. Peter (2003) Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, Vol. 50 (1), 159–175.

Zhang, Yifan (2021) PPN: The source code of the cost-sensitive portfolio policy network. GitHub. <https://github.com/Vanint/PPN>, retrieved 12.2.2021.

Zhang, Yifan – Zhao, Peilin – Li, Bin – Huang, Junzhou – Wu, Qingyao – Tan, Mingkui (2020) Cost-sensitive portfolio selection via deep reinforcement learning. Accepted Manuscript. IEEE Transactions on Knowledge and Data Engineering. <https://arxiv.org/pdf/2003.03051.pdf>, retrieved 30.1.2021.

Zimmermann, Heinz (1996) Constant return participating (CRP) portfolio insurance strategies. *Journal of Derivatives*, Vol. 4 (2), 80–88.

Zumbach, Gilles (2004) Volatility processes and volatility forecast with long memory. *Quantitative Finance*, Vol. 4 (1), 70–86.