Essays in
Asset Management and Risk Management

Universität Hamburg
Fakultät Wirtschafts- und Sozialwissenschaften
Dissertation
Zur Erlangung der Würde des Doktors der
Wirtschafts- und Sozialwissenschaften

vorgelegt von
Diplom-Wirtschaftsmathematiker
Martin Philipp Wambach
aus Osnabrück

Hamburg 2013

Vorsitzender: Prof. Dr. Michel Clement
Erstgutachter: Prof. Dr. Wolfgang Drobetz
Zweitgutachter: Prof. Dr. Alexander Szimayer

Datum der Disputation: 09.07.2013

Contents

List of Tables

List of Figures

Danksagung

Ich möchte mich bei allen Personen bedanken, die mich auf dem Weg zur Promotion am Lehrstuhl für Schiffs- und Unternehmensfinanzierung des Instituts für Finanzierung der Universität Hamburg unterstützt haben.

Insbesondere bedanke ich mich bei meinem Doktorvater, Herrn Professor Dr. Wolfgang Drobetz, der mir zu jeder Zeit sowohl auf fachlicher als auch auf persönlicher Ebene ein sehr guter Ansprechpartner gewesen ist. Er hat mir die Freiheit gegeben, eigenen Ideen nachzugehen und diese mit ihm ausführlich zu diskutieren. Sowohl die Teilnahme an nationalen und internationalen Konferenzen als auch dieser stetige Gedankenaustausch haben maßgeblich zum Gelingen und Erfolg dieser Arbeit beigetragen. Ebenso bedanke ich mich bei meinen Koautoren Hubert Dichtl und Tim Richter für die sehr angenehme Arbeitsatmosphäre sowie die zahlreichen Gespräche und wichtigen Diskussionen.

Neben der Forschung galt es auch die Lehre am Lehrstuhl zu organisieren und durchzuführen. Ich danke meinen Kollegen Dirk Schilling, Jörg Seidel, Rebekka Haller, Henning Schröder und Felix von Meyerinck, dass dies hervorragend funktioniert hat und ich nicht zuletzt durch diese Kollegen sehr gern am Lehrstuhl gearbeitet habe. An dieser Stelle möchte ich auch nicht unerwähnt lassen, dass mich insbesondere Henning Schröder zum Ende der Doktorarbeit tatkräftig unterstützt hat. Des Weiteren hat unsere Sekretärin Bettina Kourieh immer ein offenes Ohr für Probleme jeglicher Art gehabt, wodurch Sie maßgeblich zu der sehr guten Arbeitsatmosphäre am Lehrstuhl beigetragen hat. Diese Zeit werde ich stets in bester Erinnerung behalten!

Großen Dank gilt ebenfalls meiner Familie und meinen Freunden: Günter Pollner; Hannelore, Mark und Oskar Vetter; Mary Jo, Marisa und Erich Waller; Sebastian Opitz; Martin Pöhlsen; Sebastian Vogel; Oliver Garbe; Christian Neumann; Tim Böger und Antonio Wehlisch. Alle genannten Personen haben einen großen Anteil daran, dass ich mich während meiner Promotion sehr wohl gefühlt habe!

Abschließend bedarf es nicht vieler Worte, dass ich diese Doktorarbeit den drei Personen widmen möchte, die mir am nächsten stehen:

Meiner Mutter, meiner Schwester, meiner Verlobten

Chapter 1

Synopsis

1.1 Motivation

THE EXPECTED RETURN AND RISK of an investment is considered to be a key factor of the decision-making process in investment practice. Risk-adjusted performance measures such as the Sharpe ratio (Sharpe (1966)), the Omega measure (Shadwick and Keating (2002)), and the Sortino ratio (Sortino and Price (1994)) help to derive reasonable investment decisions by setting both factors in an appropriate relation to each other. The major objective of any institutional investor is to optimize the risk-adjusted performance of the invested capital within the given regulatory framework.

Once the capital has been delegated to a portfolio manager, it can be managed either actively or passively. While Hendricks et al. (1993), Wermers (2000), Avramov and Wermers (2006), Kosowski et al. (2006), Cremers and Petajisto (2009), and Fama and French (2010) all document that some active managers are in fact able to outperform their benchmark on a regular basis, the empirical analyses of Carhart (1997), French (2008), Barras et al. (2010), and Busse et al. (2010) instead provide evidence that actively managed funds produce alphas which are on average either close to zero or even negative. Based on the market's equilibrium, Sharpe (1991) supports this empirical outcome by pointing out that – before costs – actively and passively managed funds must exhibit the same return on average. In a nutshell, the market cannot outperform itself. This theoretical argument further implies that, on average, passively managed funds must earn higher net returns, which can be ascribed to the higher compensation fee of active management. Although this reasoning does not contradict the empirical findings in favor of active management, it clearly states that active managers can only win at the expense of other active managers. Thus, unless the capital to be invested cannot be allocated to the best actively managed funds, it seems well-grounded to adopt a cost-efficient, rule-based investment strategy that best reflects an investor's risk and return preference as well as his regulatory environment.

This doctoral thesis analyzes rebalancing as well as portfolio insurance strategies as cost-efficient, rule-based investment strategies by reporting statistical significance. Moreover, as risk constitutes an integral factor of the investment decision making-process, the time-varying risk of the shipping market is investigated by way of example as it represents one of the riskiest industries worldwide.

1.2 Value Added of Rebalancing

The first article "Testing Rebalancing Strategies for Stock-Bond Portfolios: Where Is the Value Added of Rebalancing?" focuses on the question whether rebalancing is able to generate a value added for institutional investors. Although this issue is of considerable importance to investment practice, it cannot be answered without further ado. On the one hand, a regular reallocation to the predefined target weights seems to be reasonable in order to satisfy the institutional investor's initially evaluated risk and return preference. On the other hand, it must not be ignored that all rebalancing strategies require the selling of a fraction of the better-performing assets and investing the proceeds in the less-performing assets. Thus, rebalancing constitutes a contrarian investment strategy.

Starting point is a two-asset class portfolio with an initial asset allocation of 60% stocks and 40% government bonds. Representing one of the world's largest institutional investors by the end of 2011, the Norway's Government Pension Fund Global (GPFG) can be cited as a prominent example of having pursued rebalancing with a 60/40 asset allocation in the past (Chambers et al. (2012)). Ranging from January 1980 to December 2011, the sample period covers 30 years with 360 monthly return observations. In contrast to all other rebalancing studies, this article does not only focus on institutional investors of the United States, but also on those of the United Kingdom and of Germany in order to examine whether country-specific characteristics could have an impact on the performance of rebalancing. Moreover, the analysis comprises investment horizons of 5, 7, and 10 years, thereby quoting realistic transaction costs of 15 bps per round-trip.

Overall, ten rebalancing strategies are under investigation which can be categorized in four distinctive rebalancing classes: (i) buy-and-hold, (ii) periodic rebalancing, (iii) threshold rebalancing, and (iv) range rebalancing. Once the capital is invested, a buy-and-hold investor holds all positions until divestment. Regardless of any market movements, no transactions take place during the investment period. In contrast, periodic rebalancing is characterized by a reallocation to the original target weights at the end of each predetermined period. In this study, the rebalancing occurs on either a yearly, a quarterly, or on a monthly basis. In addition, the construction of a no-trade interval around the target weights helps to reduce portfolio turnover and save transaction costs as an immediate consequence. While threshold rebalancing requires a reallocation to the target weights if the stocks' portfolio weight exceeds the no-trade region at the end of the period, the portfolio manager has to rebalance the assets back to the nearest edge of the no-trade region in case of range rebalancing. In line with the GPFG, a symmetric no-trade region of ±3% is implemented around the target weights.

The methodological approach builds on the stationary bootstrap of Politis and Romano (1994), which is applicable to weakly dependent, stationary data. This history-based simulation set-up helps to model realistic market conditions which are necessary in order to derive reasonable recommendations to investment practice. Most important, by drawing blocks of different lengths with replacement, time series characteristics such as short-term momentum, fat tails, and left-skewed return distributions can be preserved to the greatest possible extent. For each simulated return path, the performance measure of interest is calculated. The construction of the corresponding confidence intervals is based on Efron and Tibshirani (1998).

The analysis starts by comparing the returns of rebalancing with those of a buy-and-hold strategy and proceeds with analyzing the risk of these strategies. Taking both the return and the corresponding risk of each strategy into account, the risk-adjusted performance is evaluated with the help of the Sharpe ratio (Sharpe (1966)), the Omega measure (Shadwick and Keating (2002)), and the Sortino ratio (Sortino and Price (1994)) in a final step.

Overall, this study makes two important contributions to the academic literature. The main contribution refers to the history-based simulation set-up which allows a statistical comparison between buy-and-hold and rebalancing by reporting statistical significance levels. Secondly, while prior literature lacks a focus on institutional investors outside the United States, this article also sheds light on the issue whether country-specific characteristics affect the performance of rebalancing by extending the analysis to the financial markets of the United Kingdom and of Germany, respectively.

This study provides strong evidence that rebalancing outperforms buy-and-hold on a risk-adjusted basis. Without any exception, this finding is valid at the 1% level for all three countries under investigation (United States, United Kingdom, and Germany), all three risk-adjusted performance measures (Sharpe ratio, Sortino ratio, and Omega measure), all investment horizons (5, 7, and 10 years), and all rebalancing strategies under investigation (periodic, threshold, and range rebalancing with yearly, quarterly, and monthly trading frequencies). Moreover, while no statistical differences in returns can be detected between rebalancing and buy-and-hold, this outperformance is attributable to a reduction of portfolio risk as a consequence of the regular reallocation to the predefined target weights. In conclusion, this paper strongly supports the hypothesis that it is a risk management argument which justifies the widespread use of rebalancing in investment practice.

1.3 Optimal Rebalancing

Once rebalancing has been derived from the investor's risk and return preferences as well as his regulatory environment as an appropriate investment strategy, the resulting question of interest is which particular rebalancing strategy should be adopted. The article "Testing Rebalancing Strategies for Stock-Bond Portfolios: What Is the Optimal Rebalancing Strategy?" sheds light on this issue.

For example, with over 550 billion assets under management (AuM) by the end of December 2011 and over 19 billion AuM by the end of June 2012, respectively, both the GPFG and the Yale Endowment conduct rebalancing as a cost-efficient, rule-based investment strategy (Norwegian Ministry of Finance (2012), The Yale Endowment (2012)). With regard to these portfolio volumes, even marginal differences in risk-adjusted portfolio performance could be crucial, thereby illustrating its importance to investment practice.

Using the same set-up as before, the empirical results provide strong evidence that both excessive rebalancing (monthly periodic rebalancing) as well as overly infrequent rebalancing (yearly range rebalancing) lead to an inferior risk-adjusted portfolio performance. Although statistical significance is less pronounced for comparisons between rebalancing strategies with a similar trading interval (yearly, quarterly, or monthly) or a similar rebalancing algorithm (periodic, threshold, or range rebalancing), the results suggest that quarterly periodic rebalancing seems to provide the highest risk-adjusted portfolio performance. Moreover, time series characteristics – especially short-term momentum – tend to be the predominant driving force contributing to explain the optimal trading pattern.

1.4 Rebalancing Across Different Asset Allocations

So far, a fixed asset allocation of 60% stocks and 40% government bonds has been analyzed in order to make the results comparable with academic research as well as the investment practice. A still open question is whether a portfolio's asset allocation has an impact on the statistical significance between comparisons of rebalancing and buy-and-hold. In other words, does rebalancing provide a value added to institutional investors regardless of the asset allocation mix?

Applying the same set-up as described before, the study "Testing Rebalancing Strategies for Stock-Bond Portfolios Across Different Asset Allocations" makes two contributions to the academic literature and the investment practice. First of all, depending on the underlying country characteristics, rebalancing provides a value added if the stocks' portfolio weight exceeds a certain threshold which is approximately 30% with respect to the financial markets of the United States and the United Kingdom and about 20% regarding the financial market of Germany. However, the optimal asset allocation strongly depends on the country as well as on the period under investigation.

1.5 Portfolio Insurance

Although rebalancing leads to a significant reduction of risk, many institutional investors are not able to adopt a rebalancing strategy due to regulatory requirements. In particular, insurance companies, pension funds, and endowments have a substantial interest in preserving the invested capital. In contrast, rebalancing requires buying (selling) stocks when stocks have decreased (increased) and hence represents the sale of portfolio insurance. For this reason, the study "A Bootstrap-Based Comparison of Portfolio Insurance Strategies" analyzes different portfolio insurance strategies on a statistical basis.

The primary objective of portfolio insurance is the reduction of a portfolio's downside risk while simultaneously keeping most of its upside potential. If portfolio insurance constitutes an appropriate investment strategy that satisfies an investor's requirement, the question of which specific portfolio insurance strategy should be implemented arises. Although prior research evaluates portfolio insurance in terms of downside protection and return potential, recommendations to portfolio management differ with respect to the applied methodology, period under investigation, and market environment. For example, while the applied Monte Carlo simulation of Benninga (1990) provides evidence that a simple stop-loss rule outperforms both the constant proportion portfolio insurance strategy (CPPI) and the synthetic put strategy in terms of their terminal wealth and Sharpe ratio, the simulation results of Cesari and Cremonini (2003) indicate that CPPI and option-based portfolio insurance strategies tend to lead to superior results if information about the state of the market is neglected.

This article expands earlier studies with a focus on portfolio insurance by making two contributions to academic literature: Firstly, the bootstrap analysis proposed by Politis and Romano (1994) allows a systematic comparison between different pairs of portfolio insurance strategies by reporting statistical significance levels. In particular, the stop-loss strategy, the synthetic put strategy, the CPPI strategy, the time invariant portfolio protection strategy, and the VaR-based protection strategy are investigated by comparing the following distinguishing characteristics: (i) static versus dynamic protection; (ii) initial wealth versus accumulated wealth protection; (iii) model-based versus model-free protection; and (iv) strong floor compliance versus probabilistic floor compliance. Secondly and in contrast to Jiang et al. (2009), estimation risk is fully incorporated into the analysis as both the synthetic put strategy and the dynamic VaR-strategy require parameter forecasts of the return and its corresponding volatility. In line with Bertrand and Prigent (2011), the Omega measure is used as an appropriate risk-adjusted performance measure in order to evaluate portfolio insurance.

The simulation results document an inverse relationship between the protection quality and the return potential of the different portfolio insurance strategies. The better the downside protection is, the lower the upside potential will be. Evaluating portfolio insurance with the help of the Omega ratio, the superiority seems to be subject from the investor's required rate of return: The higher the threshold return, the more attractive insurance strategies with a higher upside potential become. In comparison to the remaining portfolio insurance strategies, the CPPI strategy exhibits a superior ratio of realized gains to losses in most instances. Finally, the superiority of the dynamic VaR-strategy strongly depends on the forecasting quality of the equity risk premium and the corresponding stock market volatility. If estimation risk is taken into account, our findings contradict those of Jiang et al. (2009) by showing no value added of the dynamic VaR-strategy.

1.6 Time-Varying Risk of Global Shipping Markets

Risk constitutes an integral component of the decision-making process in investment practice. The more pronounced an industry's risk, the higher the demand for quantifying and modeling the relevant risk sources. By way of example, the article "Dynamics of Time-Varying Volatility in the Dry Bulk and Tanker Freight Markets" concentrates on time-varying risk of the global shipping market, which has to be considered as one of the riskiest industries worldwide. For instance, featuring an annual standard deviation of almost 40% from March 1999 to October 2011, the annual volatility of changes of the Baltic Capesize Index was more than twice as high as the volatility of the MSCI World Stock Market Index. The Baltic Panamax Index even fell by more than 95% from 11,425 to 440 index points between May 2008 and December 2008. Therefore, it is of utmost importance for ship owners, operators, and shipping banks alike to have a thorough understanding of the properties of freight rate volatility in order to infer value-enhancing investment and risk management decisions.

As Kavussanos (1996a, 1996b, 1997, 2003), Glen and Martin (1998), Alizadeh and Kavussanos (2002), Chen and Wang (2004), Hui et al. (2008), Alizadeh, Nomikos, and Dellen (2011), Alizadeh and Nomikos (2009, 2011), and Roumpis and Syriopoulos (2009) all have shown that GARCH models are suitable for capturing the observed volatility clustering of freight rates, this study focuses on volatility estimates based on the class of GARCH models. While Kavussanos (1996a, 1997) as well as Alizadeh and Nomikos (2011) provide strong evidence that macroeconomic factors exhibit a significant impact on time-varying freight rates risk, Chen and Wang (2004) as well as Hui et al. (2008) further document that asymmetric effects also must not be ignored.

The resulting question of interest is whether (i) shocks from macroeconomic factors or (ii) asymmetric effects are better suited for modeling the conditional volatility (or volatility clustering) of freight rates; or (iii) whether both effects should be considered simultaneously. The more appropriate the volatility estimates are, the better freight rate risk can be managed. In order to improve freight rate risk management, it is the primary objective to find the most suitable specification that is able to capture the time-variation in the volatility of freight rates.

Focusing on the volatility structure of freight rates in the dry bulk as well as in the tanker freight markets, this analysis makes two important contributions: First of all, the impact of macroeconomic shocks and asymmetric effects on the conditional (or instantaneous) volatility of freight rates is analyzed both separately and simultaneously. Secondly and in contrast to all other studies, not only normally distributed but also t-distributed error terms are used in order to better account for fat tails.

Three important conclusions can be derived from this analysis: (i) All model specifications indicate that the assumption of a t-distribution of the error term is better suited to explain the conditional volatility (or volatility clustering) than a normal distribution. Assuming a normal distribution (as prior studies have done), some models are even misspecified. (ii) The analysis suggests that macroeconomic factors should be included in the conditional variance equation, not in the conditional mean equation. Furthermore, the number of macroeconomic factors that exhibit explanatory power decreases under a t-distribution. While the TED spread is highly significant when included in the conditional variance equation of the Baltic Dirty Tanker Index, the yield curve seems to have some explanatory power for the volatility of the Baltic Panamax Index, Baltic Capesize Index, and the Baltic Clean Tanker Index. (iii) In contrast to prior studies, asymmetric volatility effects cannot be detected in the dry bulk freight market. However, these effects are strongly pronounced in the tanker freight market.

Overall, the empirical findings have important implications for freight rate risk management. By using EGARCH-X models, a 'point-in-time measure' can be derived that accurately reflects the current risk perception of the shipping market. Both the extremely high significance levels and the values of the Akaike Information Criterion indicate that the application of the EGARCH-X model with t-distributed error terms will in all likelihood help to explain the conditional volatility of freight rates more accurately and, as a consequence, help to manage freight rate risk more efficiently.

References

Alizadeh, A.H. and M.G. Kavussanos (2002). "Seasonality Patterns in Tanker Spot Freight Rate Markets." In: *Economic Modelling* 19, pp. 747–782.

Alizadeh, A.H. and N.K. Nomikos (2009). *Shipping Derivatives and Risk Management*. Hampshire: Palgrave Macmillan.

– (2011). "Dynamics of the Term Structure and Volatility of Shipping Freight Rates." In: *Journal of Transportation and Policy* 45, pp. 105–128.

Alizadeh, A.H., N.K. Nomikos, and S. van Dellen (2011). "An Investigation into the Correct Specification for Volatility in the Shipping Freight Rate Markets." Working paper, Cass Business School.

Avramov, D. and R. Wermers (2006). "Investing in Mutual Funds When Returns Are Predictable." In: *Journal of Financial Economics* 81, pp. 339–377.

Barras, L., O. Scaillet, and R. Wermers (2010). “False Discoveries in Mutual Fund Performance: Measuring Luck in Estimated Alphas.” In: *Journal of Finance* 65, pp. 179–216.

Benninga, S. (1990). “Comparing Portfolio Insurance Strategies.” In: *Financial Markets and Portfolio Management* 4, pp. 20–30.

Bertrand, P. and J.-L. Prigent (2011). “Omega Performance Measure and Portfolio Insurance.” In: *Journal of Banking and Finance* 35, pp. 1811–1823.

Busse, J. A., A. Goyal, and S. Wahal (2010). “Performance and Persistence in Institutional Investment Management.” In: *Journal of Finance* 65, pp. 765–790.

Carhart, M.M. (1997). “On Persistence in Mutual Fund Performance.” In: *Journal of Finance* 52, pp. 57–82.

Cesari, R. and D. Cremonini (2003). “Benchmarking, Portfolio Insurance and Technical Analysis: A Monte Carlo Comparison of Dynamic Strategies of Asset Allocation.” In: *Journal of Economic Dynamics and Control* 27, pp. 987–1011.

Chambers, D., E. Dimson, and A. Ilmanen (2012). “The Norway Model.” In: *Journal of Portfolio Management* 38, pp. 67–81.

Chen, Y.-S. and S.-T. Wang (2004). “The Empirical Evidence of the Leverage Effect on Volatility in International Bulk Shipping Market.” In: *Maritime Policy & Management* 31, pp. 109–124.

Cremers, K.J.M. and A. Petajisto (2009). “How Active Is Your Fund Manager? A New Measure That Predicts Performance.” In: *Review of Financial Studies* 22, pp. 3329–3365.

Efron, B. and R.J. Tibshirani (1998). *An Introduction to the Bootstrap*. Chapman & Hall/CRC.

Fama, E.F. and K.R. French (2010). “Luck versus Skill in the Cross-Section of Mutual Fund Returns.” In: *Journal of Finance* 65, pp. 1915–1947.

French, K.R. (2008). “Presidential Address: The Cost of Active Investing.” In: *Journal of Finance* 63, pp. 1537–1573.

Glen, D.R. and B.T. Martin (1998). “Conditional Modelling of Tanker Market Risk Using Route Specific Freight Rates.” In: *Maritime Policy & Management* 25, pp. 117–128.

Hendricks, D., J. Patel, and R. Zeckhauser (1993). “Hot Hands in Mutual Funds: Short-Run Persistence of Performance, 1974-1988.” In: *Journal of Finance* 48, pp. 93–130.

Hui, W., L. Jing, and P.B. Marlow (2008). “An Analysis of Freight Rate Volatility in Dry Bulk Shipping Markets.” In: *Maritime Policy & Management* 35, pp. 237–251.

Jiang, C., Y. Ma, and Y. An (2009). “The Effectiveness of the VaR-based Portfolio Insurance Strategy: An Empirical Analysis.” In: *International Review of Financial Analysis* 18, pp. 185–197.

Kavussanos, M.G. (1996a). “Price Risk Modelling of Different Size Vessels in the Tanker Industry Using Autoregressive Conditional Heteroskedasticity (ARCH) Models.” In: *The Logistics and Transportation Review* 32, pp. 161–176.

– (1996b). “Comparisons of Volatility in the Dry-Cargo Ship Sector.” In: *Journal of Transport Economics and Policy* 30, pp. 67–82.

– (1997). “The Dynamics of Time-Varying Volatilities in Different Size Second-Hand Ship Prices of the Dry-Cargo Sector.” In: *Applied Economics* 29, pp. 433–443.

– (2003). “Time Varying Risks among Segments of the Tanker Freight Markets.” In: *Maritime Economics & Logistics* 5, pp. 227–250.

Kosowski, R., A. Timmermann, R. Wermers, and H. White (2006). “Can Mutual Fund “Stars” Really Pick Stocks? New Evidence from a Bootstrap Analysis.” In: *Journal of Finance* 61, pp. 2551–2595.

Norwegian Ministry of Finance (2012). *Report No. 17 to the Storting: The Management of the Government Pension Fund in 2011*. URL: http://www.regjeringen.no/pages/37635499/stm_gpf.pdf.

Politis, D.N. and J.P. Romano (1994). “The Stationary Bootstrap.” In: *Journal of the American Statistical Association* 89, pp. 1303–1313.

Roumpis, E. and T. Syriopoulos (2009). “Asset Allocation and Value at Risk in Shipping Equity Portfolios.” In: *Maritime Policy & Management* 36, pp. 57–78.

Shadwick, W.F. and C. Keating (2002). “A Universal Performance Measure.” In: *Journal of Performance Measurement* 6, pp. 59–84.

Sharpe, W.F. (1966). “Mutual Fund Performance.” In: *Journal of Business* 39, pp. 119–138.

– (1991). “The Arithmetic of Active Management.” In: *Financial Analysts Journal* 47, pp. 7–9.

Sortino, F.A. and L.N. Price (1994). “Performance Measurement in a Downside Risk Framework.” In: *Journal of Investing* 3, pp. 59–64.

The Yale Endowment (2012). *Endowment Update 2012*. URL: http://investments.yale.edu/images/documents/yale_endowment_12.pdf.

Wermers, R. (2000). “Mutual Fund Performance: An Empirical Decomposition into Stock-Picking Talent, Style, Transaction Costs, and Expenses.” In: *Journal of Finance* 55, pp. 1655–1695.

Chapter 2

Testing Rebalancing Strategies for Stock-Bond Portfolios: Where Is the Value Added of Rebalancing?

with H. Dichtl and W. Drobetz

Acknowledgements

We appreciate helpful discussions with Martijn Cremers, Alexander Kempf, Lawrence Kryzanowski, Philipp Kurmann, Christoph Meinerding, Rainer Schlittgen, and Heinz Zimmermann as well as with participants at the 2011 joint conference of the German Classification Society (GfKl) and the German Association for Pattern Recognition (DAGM) in Frankfurt am Main, the 2012 Midwest Finance Association Annual Meeting in New Orleans, the 2012 European Financial Management Symposium on Asset Management in Hamburg, and the 2012 Kölner Finanzmarktkolloquium on Asset Management in Cologne.

Abstract

We apply a stationary bootstrap approach that enables us to test the value added of rebalancing for stock-bond portfolios using historical data from the United States, the United Kingdom, and Germany. Analyzing the Sharpe ratio, the Omega measure, and the Sortino ratio as simple measures of value added, our history-based simulation results provide strong evidence that all rebalancing strategies significantly outperform a buy-and-hold strategy. This outperformance is attributable to reduced risk, while there are no statistical differences in returns between rebalancing and buy-and-hold. Therefore, it is a risk management argument which justifies the widespread use of rebalancing in investment practice.

2.1 Introduction

The active-passive debate has occupied both academics and practitioners for years. The central questions of interest are whether active managers can outperform their benchmark and – if so – whether this outperformance is persistent. Hendricks et al. (1993), Wermers (2000), Avramov and Wermers (2006), Kosowski et al. (2006), Cremers and Petajisto (2009), and Fama and French (2010) all provide evidence that some active managers are indeed able to generate a steady outperformance net of fees. In contrast, Carhart (1997), French (2008), Barras et al. (2010), and Busse et al. (2010) report that, on average, alphas of actively managed funds are either close to zero or even negative. Sharpe (1991) contributes to explain this empirical finding by referring to the market's equilibrium: Before costs, actively and passively managed funds must exhibit the same return, on average, as the market cannot outperform itself. However, because actively managed funds demand a higher compensation fee for their operations, it also follows that passively managed funds earn higher net returns, on average. This explanation does not contradict the empirical findings of Hendricks et al. (1993), Wermers (2000), Avramov and Wermers (2006), Kosowski et al. (2006), Cremers and Petajisto (2009), and Fama and French (2010), but it clearly demonstrates that active managers can only win at the expense of other active managers.[1] Therefore, unless the money cannot be allocated to the best actively managed funds, one should invest the money as cost-efficiently as possible by tracking an index or an investment portfolio.

Having identified an investor's risk and return preference as well as his regulatory environment, the simplest passive investment strategy a portfolio manager could implement would be buy-and-hold. However, as different assets generate different rates of

[1] For example, Cremers and Petajisto (2009) show that truly active funds can also outperform at the expense of closet index funds.

returns, a portfolio's relative asset composition will deviate from the target weights over time. In order to remain consistent with the institutional investor's initially evaluated risk and return preferences, the portfolio manager needs to rebalance the assets back to their target weights, as defined by the strategic asset allocation (Campbell and Viceira (2002)).[2] Most important, as rebalancing strategies are contrarian strategies, whereby a fraction of the better performing assets is sold and the proceeds are invested in the weaker performing assets, it is a highly challenging question whether rebalancing generates a value added for institutional investors and – if so – what the sources of this value added are.

Due to its high importance for institutional portfolio management, several aspects of rebalancing and its practical implications have been analyzed in previous studies. In order to stay focused on our own analysis of the value added of rebalancing, the following discussion concentrates on the primary research objectives and main results of rebalancing studies that are closely related to our investigation. Our empirical analysis is based on the theoretical foundations of Perold and Sharpe (1988), who discuss various portfolio strategies under different market scenarios. Focusing on a two-asset portfolio consisting of stocks and bills, they document that a buy-and-hold strategy offers a downside protection that is proportional to the amount allocated into bills, while the upside potential is proportional to the amount allocated into stocks. In contrast to buy-and-hold, rebalancing strategies exhibit less downside protection. As rebalancing requires buying stocks and selling bonds when stocks have decreased (relative to bonds), this portfolio strategy represents the sale of portfolio insurance. Moreover, facing a persistent market upswing, a frequent reallocation to the weaker performing asset also leads to a lower upside potential.[3] In contrast, rebalancing strategies perform best in relatively trendless but volatile markets, gaining advantage of the much more pronounced mean-reversion in this environment.[4] These reversals

[2] Sharpe (2010) emphasizes that, in equilibrium, not all investors can be contrarians and follow a rebalancing strategy. Rebalancing policies are adopted to reflect an investor's preferences at the time when the policy is initiated. However, over time and with changing market values of asset classes, the original asset allocation may no longer be appropriate even if an investor's characteristics are unchanged. Therefore, Sharpe (2010) proposes an adaptive asset allocation policy, where the policy proportions are adjusted over time as market values change. This strategy is macro-consistent in the sense that all investors can follow it.

[3] Leading to a concave payoff structure as shown by Perold and Sharpe (1988), Ingersoll et al. (2007) conjecture that concave payoff strategies have by construction higher Sharpe ratios and Sortino ratios than benchmark or buy-and-hold investments. However, our analysis also provides evidence that rebalancing outperforms buy-and-hold in terms of Omega ratios that incorporate all moments by considering the entire return distribution.

[4] As a result, Sharpe (2010) argues that the performance of rebalancing relative to buy-and-hold can be highly period-dependent. Irrespective of having superior knowledge about the return-generating process, investors should follow a rebalancing strategy only if they are less concerned than the average investor is about inferior returns in very bad or very good markets. Following an adaptive asset allocation policy, investors should routinely compare their asset allocations with current market

could improve portfolio returns while simultaneously reducing the risk of rebalancing strategies.

Providing empirical evidence that rebalancing strategies are indeed able to generate a value added for institutional investors, Arnott and Lovell (1993) investigate a 24-year sample period from 1968 to 1991. They infer from their analysis that rebalancing offers enhanced returns without increasing risk and recommend a monthly rebalancing strategy to investors with a long investment horizon. Examining the period from 1995 to 2004, Harjoto and Jones (2006) report that a rebalancing strategy with an incorporated no-trade interval of ±15% leads to both the highest average return and the lowest standard deviation, which in turn also results in the highest Sharpe ratio. Taken as a whole, they conclude that investors should readjust their portfolio structure, although not too frequently. Tokat and Wicas (2007) conduct Monte Carlo simulations in order to provide evidence that rebalancing is a powerful instrument for controlling risk.[5] Investigating the impact of both different market scenarios and of several rebalancing strategies, they argue that rebalancing helps to minimize risk relative to a predefined asset allocation in all market environments. This result is in line with the findings of Jaconetti et al. (2010).

Despite many similarities, our analysis of the value added of portfolio rebalancing differs from the studies presented above. In particular, we make two major contributions to the literature. Our first and paramount contribution refers to the applied methodology, which builds on the stationary bootstrap of Politis and Romano (1994). In contrast to all previous studies, we are able to conduct a systematic analysis of the value added of rebalancing by reporting statistical significance levels for the different rebalancing strategies' performance measures. We are further in the position of being able to investigate whether the value added of rebalancing arises due to a return effect, a risk effect, or both.

Our second contribution relates to the observation that prior rebalancing studies mostly focus on the US market. While Buetow et al. (2002), Masters (2003) as well as McLellan et al. (2009) consider international equities in a multi-asset class portfolio,

---

proportions in order to make sure that any differences are consistent with differences between their own circumstances and those of the average investor. Kimball et al. (2011) develop an overlapping generations model in order to illustrate how optimizing agents rebalance in equilibrium. The aggregate risk tolerance effect is a driving force in their model. Shocks to the valuation of risky assets affect investors' wealth differently, depending on their initial asset allocation. As a result, such shifts affect the distribution of wealth and change aggregate risk tolerance (violating the simple Merton (1971) model market-clearing behavior) as well as the demand for risky assets (departing from the standard model's rebalancing advice).

[5] The calibration of the mean, the volatility, and the cross-correlation parameters is based on a historical sample of the US bond and stock market from 1960 to 2003. In order to model the return-generating process of both the bond and the stock market, Tokat and Wicas (2007) assume a normal return distribution. However, Mandelbrot (1963), Fama (1965), and Clark (1973) all provide strong evidence that at least stock market returns are non-normally distributed.

Plaxco and Arnott (2002) analyze an internationally balanced portfolio consisting of bonds and stocks from 11 countries. Nevertheless, despite the fact that country-specific characteristics potentially have an impact on the implementation of the optimal rebalancing strategy, there are no studies that explicitly investigate rebalancing strategies with a focus on institutional investors outside the US. Therefore, we analyze the value added of rebalancing strategies by considering the different stock and bond market characteristics of the United States, the United Kingdom, and Germany. Overall, these two contributions – deriving statistical inference and using an international data set – constitute the novel path that our analysis takes and which separates it from previous rebalancing studies.

Our findings have immediate practical implications. First of all, despite the strong performance of stocks relative to bonds during the 30-year long sample period, our simulation results do not uncover statistical evidence that the average return of a buy-and-hold strategy is, on average, higher than that of different rebalancing strategies. Secondly, the average risk of rebalancing is significantly reduced when compared with buy-and-hold. Thirdly, all rebalancing strategies significantly outperform buy-and-hold in terms of average risk-adjusted performance measured by the Sharpe ratio, the Omega measure, and the Sortino ratio. Our empirical findings are robust for all three countries under investigation, for all analyzed trading frequencies, and for all investment horizons, thus contributing to the explanation as to why rebalancing strategies are popular in investment practice.

The remainder of this paper is structured as follows: Section 2.2 briefly describes the data and provides descriptive statistics, while Section 2.3 classifies the implemented rebalancing strategies. The test design of the stationary bootstrap approach is outlined in Section 2.4. Section 2.5 presents and discusses the results of our history-based simulation analysis. Section 2.6 reports on various robustness checks. The paper concludes in Section 2.7, where implications for portfolio management and institutional investors are pointed out.

2.2 Data Description

2.2.1 Dataset

As country-specific characteristics could have an impact on rebalancing, we not only concentrate on domestic institutional investors in the United States, but also on those in the United Kingdom and Germany. We use monthly return data of well-diversified stock and government bond market total return indices as well as money market rates for each country from Thomson Datastream. The sample period ranges from January 1982 to December 2011. This long 30-year period is necessary in order to implement a statistical test, but government bond time series of this length are only available for the financial markets of the United States, the United Kingdom, and Germany. Exhibiting maturities of 5, 7, and 10 years, these government bond time series also determine the length of the respective investment horizons to be analyzed. We use Treasury bills (United States), LIBOR (United Kingdom), and FIBOR (Germany) as proxies for the risk-free rates with 3-month maturities.

2.2.2 Descriptive Statistics

The implementation of the stationary bootstrap approach is motivated by two reasons. First of all, given the country-specific characteristics of the different financial markets, one cannot assume that particular relationships which hold in one country are also observable in any other country. Panel A of Table I illustrates these cross-sectional differences between the stock, government bond, and money markets of the United States, the United Kingdom, and Germany over the period from January 1982 to December 2011. For example, the German stock market exhibits the highest annualized volatility of all three countries featuring a value of 22.06%, whereas the German government bond market simultaneously has the lowest annualized volatility with a value of 5.53%. Thus, an analysis of the three financial markets in our sample helps to corroborate whether our empirical findings are robust in the cross-section of countries.

Secondly, prior research has already shown that the time series properties themselves can change over time, making it almost impossible to appropriately calibrate the parameters for a Monte Carlo simulation.[6] However, by using historical data, all times series information is fully incorporated into our simulation analysis. In order to get a detailed insight into the time variation of the underlying time series characteristics, we divide the entire 30-year sample period into two disjunctive 15-year subsamples. Although the time series characteristics of the United Kingdom and Germany are different than those of the United States, all three countries exhibit qualitatively similar patterns in terms of business cycles. Panel B of Table I, by way of example, shows the descriptive statistics of the US stock, government bond, and money market over the full sample period as well as the two corresponding 15-year subsamples. Clearly, substantial variation is exhibited by the distributional characteristics over time. For example, the US stock market features an average annualized return of 15.59% over the period from January 1982 to December 1996, which drops to only 5.31% over the period from January 1997 to December 2011.

[6] Cf. Gibbons and Ferson (1985), Ferson et al. (1987), Ferson and Harvey (1991), and Ferson and Harvey (1993) for studies related to time-varying risk premia; Engle (1982), Engle et al. (1987), Ng (1991), and Dumas and Solnik (1995) for research on time-varying risk, as well as Erb et al. (1994), Ball and Torous (2000), Longin and Solnik (2001), and Buraschi et al. (2010) for studies with a focus on time-varying asset class correlations.

Table I – Descriptive Statistics

Panel A presents the cross-sectional descriptive statistics of the stock, government bond, and money markets of the United States, the United Kingdom, and Germany over the entire 30-year sample period from January 1982 to December 2011. Panel B shows the descriptive statistics of the United States over the entire 30-year sample period as well as the two corresponding disjunctive 15-year subsamples. Bonds denote government bonds with a maturity of 10 years. Cash represents 3-month money market rates. All statistics are calculated on a monthly basis using continuous compounded returns. Mean, Volatility, Skewness, and Kurtosis denote the annualized mean return, volatility, skewness, and kurtosis. Minimum and Maximum are the monthly minimum and maximum returns.

Panel A: Cross-Sectional Descriptive Statistics

Asset	Statistics	United States	United Kingdom	Germany
Stocks	Mean (%)	10.45	10.84	8.75
	Volatility (%)	15.77	16.14	22.06
	Skewness	−0.91	−1.15	−0.92
	Kurtosis	6.07	8.05	5.60
	Minimum (%)	−23.85	−30.02	−28.67
	Maximum (%)	12.47	13.72	19.02
Bonds	Mean (%)	8.57	10.19	7.34
	Volatility (%)	7.91	8.01	5.53
	Skewness	0.05	−0.06	−0.29
	Kurtosis	3.66	4.45	3.26
	Minimum (%)	−7.36	−8.16	−5.69
	Maximum (%)	9.40	8.17	5.37
Cash (level)	Mean (%)	4.46	6.91	4.43
	Volatility (%)	0.77	1.01	0.65
	Skewness	0.16	0.23	0.55
	Kurtosis	2.70	2.39	2.62
	Minimum (%)	0.00	0.00	0.00
	Maximum (%)	0.01	0.01	0.01

Panel B: Descriptive Statistics of the United States for Subsamples

Asset	Statistics	Full Sample Jan-82 - Dec-11	1st Half Jan-82 - Dec-96	2nd Half Jan-97 - Dec-11
Stocks	Mean (%)	10.45	15.59	5.31
	Volatility (%)	15.77	14.47	16.89
	Skewness	−0.91	−1.12	−0.71
	Kurtosis	6.07	9.79	3.94
	Minimum (%)	−23.85	−23.85	−18.76
	Maximum (%)	12.47	12.47	10.42
Bonds	Mean (%)	8.57	10.66	6.48
	Volatility (%)	7.91	8.10	7.68
	Skewness	0.05	0.08	−0.01
	Kurtosis	3.66	2.85	4.63
	Minimum (%)	−7.36	−4.50	−7.36
	Maximum (%)	9.40	7.30	9.40
Cash (level)	Mean (%)	4.46	6.21	2.71
	Volatility (%)	0.77	0.60	0.57
	Skewness	0.16	0.45	−0.01
	Kurtosis	2.70	3.07	1.44
	Minimum (%)	0.00	0.00	0.00
	Maximum (%)	0.01	0.01	0.01

2.3 Implemented Rebalancing Strategies

The academic literature as well as institutional portfolio managers differentiate between periodic and interval rebalancing strategies. Advising a periodic rebalancing mandate, a portfolio manager has to rebalance the assets to their initial target weights at the end of each predetermined period (e.g., yearly, quarterly, or monthly). In contrast, an interval rebalancing mandate requires the portfolio manager to adjust the asset allocation whenever an asset moves beyond a prespecified threshold (e.g., ±3%, ±5%, or ±10%). Our study focuses on a mixture of both methodologies: periodic rebalancing with the additional option to incorporate a symmetric no-trade interval around the target weights.

Furthermore, it is necessary to distinguish between the two resulting approaches with regard to the implementation of the symmetric no-trade interval. In particular, when an asset exceeds the predetermined interval boundaries, either a strict adjustment to the target weights (Buetow et al. (2002), Harjoto and Jones (2006)) or a rebalancing to the corresponding interval boundaries (Leland (1999)) must be implemented. Following the argument of Perold and Sharpe (1988), who emphasize that different strategies can produce remarkable differences in risk and return characteristics, we implement the most common rebalancing strategies: (i) buy-and-hold, (ii) periodic rebalancing, (iii) periodic interval rebalancing with a strict adjustment to the initial target weights (threshold rebalancing), and (iv) periodic interval rebalancing with a reallocation to the nearest edge of the corresponding thresholds (range rebalancing). With respect to these strategies, we look at yearly, quarterly, and monthly trading frequencies. Table II presents the resulting classification of all implemented rebalancing strategies.

Table II – Classification of Implemented Rebalancing Strategies

This table presents all rebalancing strategies under investigation. The periodic rebalancing strategies 2, 3, and 4 are characterized by a regular reallocation to the predetermined target weights at the end of each period. Strategies 5, 6, and 7 represent threshold rebalancing, which is classified as periodic interval rebalancing with a strict adjustment to the target weights. In contrast, the range rebalancing strategies 8, 9, and 10 require a reallocation to the nearest edge of the predefined interval boundaries. A threshold of ±3% is applied to both threshold rebalancing and range rebalancing.

Rebalancing Strategy	Frequency	Threshold	Reallocation	Classification	No.
Buy-and-Hold	No Adjustments	No Threshold	No Reallocation	Buy-and-Hold	1
Yearly Periodic Rebalancing	Yearly	No Threshold	Target Weights	Periodic	2
Quarterly Periodic Rebalancing	Quarterly	No Threshold	Target Weights	Periodic	3
Monthly Periodic Rebalancing	Monthly	No Threshold	Target Weights	Periodic	4
Yearly Threshold Rebalancing	Yearly	Threshold	Target Weights	Threshold	5
Quarterly Threshold Rebalancing	Quarterly	Threshold	Target Weights	Threshold	6
Monthly Threshold Rebalancing	Monthly	Threshold	Target Weights	Threshold	7
Yearly Range Rebalancing	Yearly	Threshold	Interval Boundaries	Range	8
Quarterly Range Rebalancing	Quarterly	Threshold	Interval Boundaries	Range	9
Monthly Range Rebalancing	Monthly	Threshold	Interval Boundaries	Range	10

A simple example demonstrates how our periodic interval rebalancing methodology works. Assume a 60% stocks and 40% bonds asset allocation with a quarterly rebalancing frequency and a threshold of ±3% around the target weights. The portfolio strategy '3% quarterly threshold rebalancing' implies a strict adjustment to the original stock allocation of 60% whenever the stock allocation exceeds the threshold of ±3% at the end of each quarter. In contrast, the portfolio strategy '3% quarterly range rebalancing' requires the asset manager to check whether the weight of stocks exceeds 63% or falls below 57% of the portfolio's current market capitalization at the end of each quarter. In the first case, the manager must rebalance stocks to the upper threshold of 63%, whereas in the second case an adjustment of stocks to the lower threshold of 57% is required. In all other cases, no transactions are necessary because the stocks' target weight falls within the predetermined no-trade interval [57%; 63%]. According to Leland (1999), this approach reduces transaction costs and may potentially lead to superior portfolio performance. If no thresholds are specified, both threshold and range rebalancing are reduced to the more general periodic approach. In addition, if no trading intervals are determined, periodic rebalancing is reduced to buy-and-hold as the most general rebalancing strategy, which implies no rebalancing at all.

We concentrate on a two-asset-class portfolio with an initial asset allocation of 60% stocks and 40% government bonds.[7] On the one hand, this approach adequately reflects common investment behavior in practice (Chambers et al. (2012)). On the other hand, it allows the comparison of our empirical findings with related rebalancing studies. Despite our focus on only two asset classes for the purpose of simplification, one should consider that each index constitutes a well-diversified representative of an entire asset class of the corresponding country. In addition, we also model realistic transaction costs of 15 bps per round-trip in all our simulations. Particularly, we quote 10 bps for buying/selling stocks and 5 bps for selling/buying bonds.[8]

[7] In order to conduct a comprehensive analysis of whether rebalancing is able to generate a value added for institutional investors, we vary the parameters country (United States, United Kingdom, and Germany), the implemented rebalancing strategies (buy-and-hold as well as periodic, threshold, and range rebalancing), the trading frequencies (yearly, quarterly, and monthly), the performance measures (return, volatility, semi-volatility, Sharpe ratio, Omega measure, and Sortino ratio), and the investment horizons (5, 7, and 10 years). As all these determinants are linked by multiplication, we have to keep the initial asset allocation constant in order to stay focused on the main contribution of our study. However, being one of the world's largest institutional investors, Norway's Government Pension Fund Global (GPFG) is a prominent example of having pursued a 60% stocks and 40% bonds asset allocation in the past (Chambers et al. (2012)). In consideration of the GPFG's long-term strategy, these portfolio weights have been slightly modified to 60% stocks, 35% bonds, and 5% real estate. Moreover, it is worth noting that the GPFG does not follow an adaptive asset allocation (Sharpe (2010)), but instead a 3% quarterly threshold rebalancing strategy (Norwegian Ministry of Finance (2012)).

[8] The transaction costs of 15 bps per round-trip solely relate to the assets' reallocation, which in turn depends on the underlying rebalancing algorithm. Additional costs do incur with respect to the administration and the management of the portfolio. Exchange traded funds (ETFs) represent a cost-effective method to implement rule-based portfolio strategies such as rebalancing. According to

2.4 Statistical Inference

Previous research based on historical analyses remains incomplete as it merely investigates a single realization or a fairly small number of realizations of stock and bond returns. This issue is reinforced by the fact that the performance of rebalancing is highly path-dependent because it constitutes a dynamic portfolio strategy. For example, Arnott and Lovell (1993) show that rebalancing offers enhanced returns without increasing risk, whereas Jaconetti et al. (2010) provide evidence that the primary objective of rebalancing is the reduction of risk. The only systematic finding of most rebalancing studies is that a buy-and-hold strategy seems to underperform rebalancing strategies when both the return and the risk of these strategies are taken into account. But even in this case, a major concern is whether these findings are statistically significant. It is possible that the return observations are more influenced by specific characteristics of the underlying sample period rather than by the properties of the rebalancing strategy under investigation. As the danger of data snooping described can be severe (Brock et al. (1992)), the empirical results of these studies do not allow reliable interpretations. Dividing the sample period into disjunctive subsamples, e.g., up- and downswings of the stock market (Harjoto and Jones (2006)), does not solve this fundamental inference problem either, as this procedure cannot generate enough observations to conduct a statistical test.

Accordingly, many prior studies apply Monte Carlo simulations for evaluating rebalancing strategies (Jones and Stine (2010), Sun et al. (2006), Donohue and Yip (2003), Buetow et al. (2002)). Changing stock, bond, and money market characteristics and their impact on rebalancing strategies can be examined in more detail because Monte Carlo simulations allow for deriving the entire return distribution under different economic scenarios. However, this simulation technique generally suffers from the shortcoming that it is not based on real-world data. If time series characteristics of assets as well as those of entire financial markets are neither correctly nor completely incorporated, simulation results could be biased. Most important, Monte Carlo simulations often assume normally distributed stock returns even though stock market returns generally violate a normality assumption by exhibiting fat tails and heteroscedasticity as well as by tending to be left-skewed (Annaert et al. (2009)). Therefore, Eraker (2004) suggests a stochastic volatility process with jumps in asset values. This process has the geometric Brownian motion as a special case, but allows for heavier tails in the return

the market leaders iShares, Lyxor Asset Management, and db X-trackers, the total expense ratio (TER) of the most liquid ETFs ranges between 15 and 20 bps for government bonds and between 15 and 52 bps for equities. These costs are independent of the rebalancing frequency and are charged regardless of the applied portfolio strategy. Therefore, we exclude the TER from our analysis as it does not affect the issue whether rebalancing provides a value added to institutional investors.

distribution. Moreover, De Bondt and Thaler (1985), Poterba and Summers (1988), as well as Brennan et al. (2005) all provide evidence that stock returns exhibit positive autocorrelation in the short-run and mean reversion in the long-run. Finally, asset class correlations tend to increase during recession periods (Longin and Solnik (2001)). While Monte Carlo simulations are unable to capture all of these return characteristics appropriately, a statistical test that is based on historical data is more suitable for incorporating all different time series properties.

Due to these shortcomings of both historical analyses and Monte Carlo simulations, we perform simulations that are based on real-world data by implementing the stationary bootstrap of Politis and Romano (1994). This history-based simulation clearly separates our investigation of the value added of rebalancing both from historical analyses and Monte Carlo simulations. Analyzing the value added of rebalancing, we test whether the mean of a difference time series is equal to zero:

$$H_0 : \Delta_{PM} = 0 \qquad \text{versus} \qquad H_1 : \Delta_{PM} \neq 0, \tag{1}$$

where PM denotes the performance measure of interest, which is either the return, the volatility, the semi-volatility, the Sharpe ratio, the Sortino ratio, or the Omega measure. The difference between the two performance measures is given by:

$$\Delta_{PM} = PM_A - PM_B, \tag{2}$$

where A and B constitute rebalancing strategies. An appropriate point estimator of (2), which is defined as the arithmetic mean, is given by:

$$\widehat{\Delta}_{PM} = \widehat{PM}_A - \widehat{PM}_B. \tag{3}$$

In order to remain focused on our main contribution – the statistical comparison of rebalancing and buy-and-hold – we investigate the difference time series of a predetermined performance measure between any rebalancing strategy and buy-and-hold. As classified in Table II, this could either be periodic rebalancing, threshold rebalancing, or range rebalancing. Overall, we end up with three comparisons for each rebalancing class, each country, and each investment horizon:

$$\begin{aligned} &\text{Monthly rebalancing - buy-and-hold (M-BAH)}, && (4.1)\\ &\text{Quarterly rebalancing - buy-and-hold (Q-BAH)}, && (4.2)\\ &\text{Yearly rebalancing - buy-and-hold (Y-BAH)}. && (4.3) \end{aligned}$$

By bootstrapping return paths that could have been realized in the past, we are able to efficiently evaluate the available information of the underlying sample period. Although bootstrap techniques destroy the original path as there can be only one single realization, drawing random blocks of different lengths with replacement from our 30-year sample period allows us to preserve time series' properties and financial markets' dependencies (such as positive autocorrelation in the short-run, heteroscedasticity, fat tails, left-skewed return distributions, and asset class correlations) to the greatest possible extent. In particular, we implement the stationary bootstrap of Politis and Romano (1994), which is applicable to stationary, weakly dependent data.[9] The only parameter left to be specified is the probability P for resampling the return observations. As P follows a geometric distribution, we take advantage of the resulting inverse relationship between P and the average block size. Thus, it is sufficient to determine the average block size to be drawn as it constitutes the expected reciprocal value of P. Even though Ledoit and Wolf (2008) report that the stationary bootstrap of Politis and Romano (1994) is quite insensitive to the choice of the average block size, we ascertain the optimal average block length by applying the automatic block-length selection for the dependent bootstrap of Politis and White (2004) as well as the corrections made by Patton et al. (2009). Overall, an average block length of 2 is suggested for both the stock market and the government bond time series of all three countries under investigation.[10] Therefore, we are able to compare our empirical findings despite the cross-sectional differences between the stock and government bond markets of the United States, the United Kingdom, and Germany as shown in Table I.

In a first step, we resample 100 return paths of stocks, government bonds, and risk-free rates.[11] In order to retain the cross-sectional dependency structure, we conduct a pairwise resampling. The length of these return paths is determined by the investment horizon of either 5, 7, or 10 years. That is, we resample investment horizons of 5, 7, and 10 years from the original 30-year sample period.[12] Although the investment horizons to be examined do not match our 30-year sample period, this procedure allows us to analyze the impact of different investment horizons which would not be possible otherwise (when exploiting the information of the full data set). Having determined

---

[9] The implementation of the stationary bootstrap exactly follows the algorithm in Politis and Romano (1994).

[10] As a robustness check, we also test an average block size of $b \in 4,6,8$ in Section 2.6.3. The empirical results are qualitatively the same.

[11] In general, bootstrap methodologies such as the applied stationary bootstrap of Politis and Romano (1994) require stationary processes. Representing highly non-stationary processes, none of the applied money market rates fulfills this necessary requirement. By using a bootstrap approach, an investigation of the cash market's volatility would induce a volatility that could be attributed to the non-stationary characteristics of the cash market. Nevertheless, as the average money market rates are only included in the calculations of the Sharpe ratio, these difficulties do not emerge in our analysis.

[12] In contrast, Ledoit and Wolf (2008, 2011) investigate a 10-year investment horizon by resampling from a 10-year sample period.

both the rebalancing class and the performance measure of interest, we calculate the mean for each of the three possible difference time series according to (4.1) – (4.3). In a second step, we repeat this procedure B times in order to construct two-sided percentile intervals as described by Efron and Tibshirani (1998), where[13]:

$$\widehat{\Delta}^*_{PM[1]} \leq \widehat{\Delta}^*_{PM[2]} \leq \ldots \leq \widehat{\Delta}^*_{PM[B-1]} \leq \widehat{\Delta}^*_{PM[B]} \tag{5}$$

denotes the ordered difference series of the performance measure of interest. Based on this difference series, a confidence interval can be constructed as follows:

$$CI = \left[\widehat{\Delta}^*_{PM\left[\frac{\alpha}{2}\cdot B\right]}, \widehat{\Delta}^*_{PM\left[\left(1-\frac{\alpha}{2}\right)\cdot B\right]}\right]. \tag{6}$$

The null hypothesis H_0 is rejected at the significance level α if $0 \notin CI$. The nominal levels to be considered are 0.01, 0.05, and 0.10. Following Ledoit and Wolf (2008), we conduct $B = 1,000$ simulations. Repeated simulations reveal that our results (in terms of statistical significance) are stable in capturing the underlying patterns in our sample.

2.5 Empirical Simulation Results

This section presents the main results of our simulation analyses. We start our discussion by comparing the returns of a buy-and-hold strategy and periodic rebalancing with yearly, quarterly, and monthly trading intervals. We proceed by analyzing the risk of these strategies. Taking both the return and the corresponding risk of each strategy into account, we finally evaluate the risk-adjusted performance.

2.5.1 Returns

Any rebalancing strategy requires the selling of a fraction of the better performing assets and investing the proceeds in the weaker performing assets. Focusing on the portfolio return as the measure of interest, it seems reasonable to assume that buy-and-

[13] Focusing on small sample sizes, Romano and Wolf (2006) show that the studentized block bootstrap leads to an improved coverage accuracy compared with normal theory intervals as well as the basic bootstrap. Ledoit and Wolf (2008, 2011) also recommend a studentized time series bootstrap for calculating p-values, if small to moderate sample sizes are to be investigated. However, as our sample includes 360 monthly observations, it can be considered as rather large, thereby supporting the construction of percentile intervals. Besides our primary objective to compare rebalancing strategies on a statistical basis, this procedure is also straightforward to implement. Most important, with respect to the analyzed risk-adjusted performance measures Sharpe ratio, Sortino ratio, and Omega measure, our main empirical results are all significant at the 1% level. Repeated simulations show that these findings are robust. Therefore, we assume that our conclusions will not change by constructing confidence intervals with a potentially better asymptotic coverage.

hold strategies outperform rebalancing strategies with increasing investment horizons. Provided that one asset outperforms the other in every single period, this notion is always correct given the mechanics of rebalancing. However, this conclusion cannot be drawn if real-world data is to be considered.

Table III – Development of Net Asset Values: United States
Classified by investment horizon, this table illustrates the development of a $100 investment over the sample period from January 1982 to December 2011. The total amount is invested in the US market as a lump-sum payment in January 1982. All strategies are based on a 60% stocks and 40% bonds asset allocation with a threshold of 0%. Transaction costs are quoted at 15 bps per round-trip.

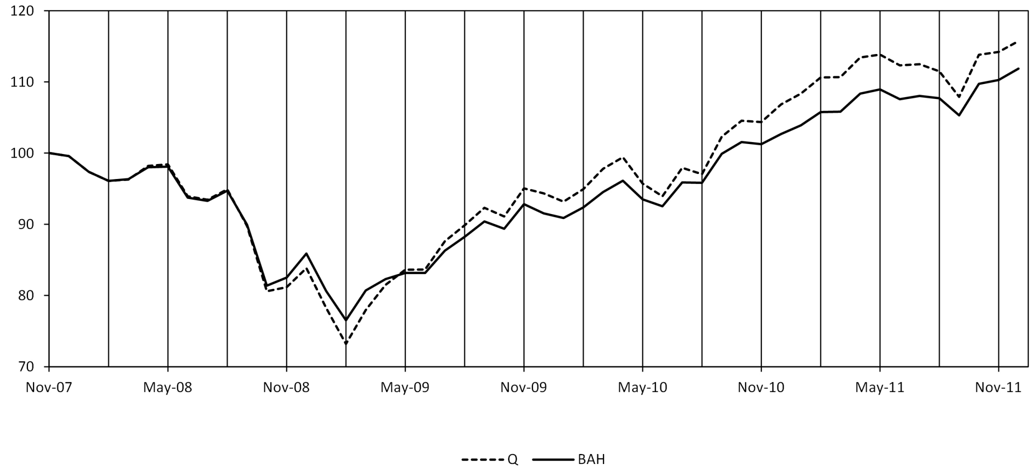
Investment Horizon	Period	Stocks	Bonds	BAH	Yearly Rebalancing	Quarterly Rebalancing	Monthly Rebalancing
5	01/82-12/86	246.6	242.6	244.9	246.1	246.6	247.2
10	01/82-12/91	501.4	364.0	446.4	445.8	455.6	454.4
15	01/82-12/96	1,036.0	494.7	819.5	785.4	799.0	796.9
20	01/82-12/01	1,710.0	677.4	1,296.9	1,237.4	1,264.8	1,247.2
25	01/82-12/06	2,303.3	853.8	1,723.5	1,669.1	1,699.6	1,664.3
30	01/82-12/11	2,298.3	1,306.8	1,901.7	2,117.2	2,121.7	2,034.6

As an example, Table III illustrates the development of a $100 investment over the sample period from January 1982 to December 2011. Although buy-and-hold exhibits the highest NAV after 15, 20, and 25 years, it is clearly dominated by all other periodic rebalancing strategies after 30 years. Thus, despite the strong performance of stocks relative to bonds during the entire sample period, one cannot necessarily conclude that buy-and-hold performs better than rebalancing. As rebalancing is a dynamic portfolio strategy, its performance is path-dependent.

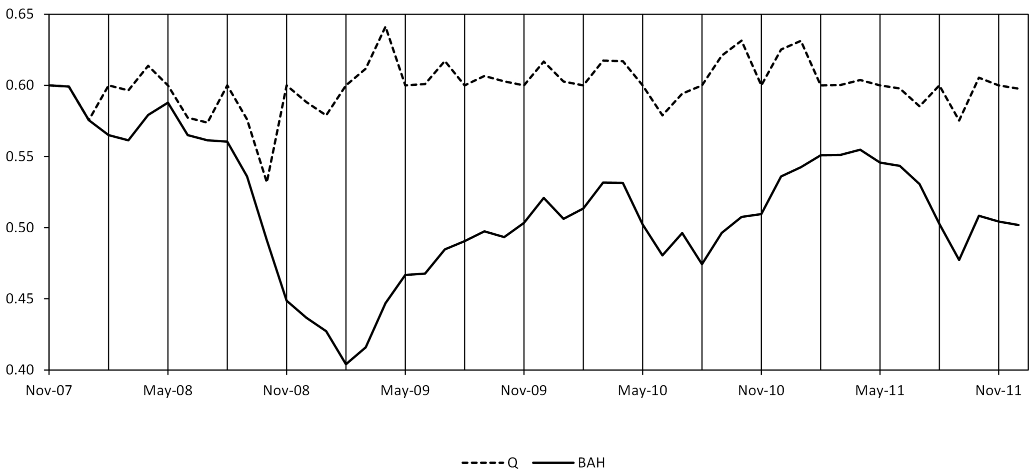
Analyzing the path dependency in more detail, Panel A in Figure I presents, by way of example, the different developments of a $100 investment of both an underlying quarterly periodic rebalancing and a buy-and-hold strategy. In order to illustrate the impact of different market scenarios on the resulting performances of these two strategies, the period under investigation covers the last recession from December 2007 to June 2009 as well as the subsequent market trends until December 2011 (National Bureau of Economic Research (2012)). Panel B depicts the corresponding relative market capitalization of stocks of both strategies at the beginning of each month after the rebalancing event has taken place. As shown in Panel A, quarterly periodic rebalancing performs worse compared with buy-and-hold during the prolonged stock market meltdown in 2008, which caused a decline of the US stock market capitalization by almost 50%. This observation is explained by the regular reallocation at the end of each quarter to the initial 60/40 asset allocation. In a trending market environment with falling stock prices, frequent rebalancing leads to inferior NAVs. Panel A further reveals that during the subsequent market upswing, quarterly rebalancing outperforms the buy-and-hold strategy. This finding can be traced back to the fact that the performance

Figure I – Performance of a $100-Investment and its Portfolio Weights of Stocks

Panel A of Figure I presents the performance of a $100-investment of both buy-and-hold (BAH) as well as quarterly periodic rebalancing (Q). Panel B plots the corresponding stock portfolio weights at the beginning of each month. In the case of quarterly periodic rebalancing, the portfolio weights of each third month are shown after the rebalancing event has taken place. Both strategies are based on a 60% stocks and 40% bonds asset allocation with a threshold of 0%. Transaction costs are quoted at 15 bps per round-trip. In order to better illustrate the mechanics of rebalancing, the subsample period starts at the beginning of the last contraction period in December 2007, as determined by the National Bureau of Economic Research (2012).



(A) Performance of a $100-Investment



(B) Portfolio Weights of Stocks

of an investment strategy not only depends on the return of the underlying assets, but also on their corresponding portfolio weights. In particular, during the following recovery, quarterly periodic rebalancing produces higher NAVs compared with buy-and-hold because of its initial 60/40 stock-bond allocation at the start of the recovery and the immediate readjustment at the end of each quarter. In contrast, buy-and-hold is disadvantaged by the decrease to a much lower stock allocation when the recovery starts. The initial stock-bond allocation at the lower turning point is roughly 40/60 (rather than 60/40) because of the poor stock performance during the prior market crash. As illustrated in Panel B, the stock allocation cannot recover from this market crash within the remaining investment period. Due to its lower average stock allocation in the subsequent upside market, buy-and-hold is outperformed by quarterly periodic rebalancing. This observation supports Perold and Sharpe's (1988) conjecture that rebalancing strategies perform best in volatile sideway markets, whereas buy-and-hold leads to superior results in strongly pronounced market upswings and downswings, respectively.

A resulting consequence of the rebalancing strategies' path dependency is that all empirical findings of rebalancing analyses are highly dependent on the period under investigation, a fact that is illustrated in Table III as well as in Figure I. Therefore, rebalancing studies that consider only one single realization are potentially exposed to the problem of data snooping, which could lead to non-optimal recommendations to portfolio management and institutional investors. Accordingly, our analysis focuses on the average performance in terms of returns, risk, and risk-adjusted performance measures. By using a 30-year historical data sample, our analysis takes changing stock, government bond, and money market conditions as well as resulting cross-sectional dependencies appropriately into account in order to derive reliable recommendations for investment practice.

Table IV – Average Annualized Returns of Periodic Rebalancing

Classified by investment horizon, this table shows the average annualized returns of buy-and-hold as well as periodic rebalancing with yearly, quarterly, and monthly trading intervals over the sample period from January 1982 to December 2011. All strategies are based on a 60% stocks and 40% bonds asset allocation with a threshold of 0%. Transaction costs are quoted at 15 bps per round-trip. 1,000 simulations with an average block length of 2 are performed. Repeated simulations reveal that the results are stable.

Period	Rebalancing Strategy	United States	United Kingdom	Germany
5	Buy-and-Hold	9.80	10.27	8.64
5	Yearly Rebalancing	9.78	10.25	8.64
5	Quarterly Rebalancing	9.76	10.24	8.60
5	Monthly Rebalancing	9.72	10.23	8.54
7	Buy-and-Hold	10.13	10.82	8.96
7	Yearly Rebalancing	10.12	10.82	8.97
7	Quarterly Rebalancing	10.10	10.81	8.93
7	Monthly Rebalancing	10.06	10.80	8.87
10	Buy-and-Hold	10.13	10.92	8.88
10	Yearly Rebalancing	10.11	10.92	8.90
10	Quarterly Rebalancing	10.09	10.91	8.87
10	Monthly Rebalancing	10.06	10.90	8.80

Classified by investment horizon and country, Table IV illustrates the average annualized returns of buy-and-hold and periodic rebalancing with yearly, quarterly, and monthly trading intervals. The results are mixed and imply that the return differences tend to be of rather marginal economic importance.

Table V – CIs: Average Annualized Returns of Periodic Rebalancing

Classified by investment horizon, this table shows the confidence intervals of the difference time series of the average annualized returns between periodic rebalancing and buy-and-hold. The sample period ranges from January 1982 to December 2011. All strategies are based on a 60% stocks and 40% bonds asset allocation with a threshold of 0%. Transaction costs are quoted at 15 bps per round-trip. BAH denotes buy-and-hold, Y yearly periodic rebalancing, Q quarterly periodic rebalancing, and M monthly periodic rebalancing. For example, M-BAH denotes the difference time series of 'Monthly periodic rebalancing minus buy-and-hold'. For each two strategies that are compared, the lower and upper boundary of the confidence interval is calculated. 1,000 simulations with an average block length of 2 are performed. Repeated simulations reveal that the results are stable. *, **, and *** represent statistical significance at the 10%, 5%, and 1% level, respectively. If there is no statistical significance, the corresponding 10% quantiles are reported.

Period	Strategy	United States		United Kingdom		Germany	
5	M-BAH	−0.0017	0.0000	−0.0012	0.0002	−0.0026	0.0005
5	Q-BAH	−0.0014	0.0003	−0.0010	0.0004	−0.0020	0.0010
5	Y-BAH	−0.0011	0.0005	−0.0009	0.0005	−0.0014	0.0013
7	M-BAH	−0.0016	0.0001	−0.0010	0.0005	−0.0026	0.0006
7	Q-BAH	−0.0012	0.0005	−0.0008	0.0007	−0.0019	0.0012
7	Y-BAH	−0.0010	0.0007	−0.0007	0.0007	−0.0015	0.0016
10	M-BAH	−0.0017	0.0001	−0.0010	0.0005	−0.0024	0.0007
10	Q-BAH	−0.0014	0.0004	−0.0008	0.0007	−0.0017	0.0013
10	Y-BAH	−0.0011	0.0006	−0.0008	0.0008	−0.0013	0.0016

Table V reports whether these return differences are statistically significant or whether they can simply be attributed to specific characteristics of the underlying sample period. If both boundaries are positive (negative), rebalancing boasts a significantly higher (lower) average annualized return compared with buy-and-hold. Otherwise, the confidence interval includes zero, implying that the difference is lost in estimation error and that no statistical inferences can be drawn. Substantiating the findings of Table IV, our analysis documents that there are no statistically significant differences in returns between buy-and-hold and periodic rebalancing. In contrast, Perold and Sharpe (1988) report that reversals in stock markets could improve portfolio returns when applying a rebalancing strategy. However, as they conduct their analysis for volatile sideway markets, our empirical results do not contradict this argument. Instead, there are two effects that work in opposite directions. As illustrated in Table III by way of example, stock markets substantially outperformed government bond markets during our entire sample period. While a high market volatility advantages rebalancing compared with buy-and-hold, buy-and-hold benefits from a well-pronounced positive market trend, which is induced by the better average performance of stocks relative to government bonds. As a result, there is no statistical significance in returns because both effects outweigh each other on average.

2.5.2 Risk

So far, our investigation has pointed out that there are no differences in average returns between buy-and-hold and periodic rebalancing. Frequent rebalancing must thus offer other key benefits that explain its importance for institutional investors. By showing that dynamic portfolio strategies, such as rebalancing, produce different risk and return characteristics, Perold and Sharpe (1988) emphasize that the choice of an appropriate strategy is subject to the investor's risk preference. Therefore, not only the return of a strategy itself, but also its risk must be carefully taken into account.

2.5.2.1 Volatility

In order to further analyze the value added of rebalancing, Table VI presents both the average annualized volatility and the average annualized semi-volatility classified by strategy, investment horizon, and country. As buy-and-hold boasts the highest average annualized volatility for all investment horizons and all countries, the empirical results in Panel A of Table VI indicate that rebalancing may consistently lead to a lower volatility.

Table VI – Average Annualized Risk of Periodic Rebalancing
Classified by investment horizon, this table shows both the average annualized volatilities and the average annualized semi-volatilities of buy-and-hold as well as periodic rebalancing with yearly, quarterly, and monthly trading intervals. The sample period ranges from January 1982 to December 2011. All strategies are based on a 60% stocks and 40% bonds asset allocation with a threshold of 0%. Transaction costs are quoted at 15 bps per round-trip. 1,000 simulations with an average block length of 2 are performed. Repeated simulations reveal that the results are stable.

Period	Rebalancing Strategy	United States	United Kingdom	Germany
	Panel A: Average Annualized Volatility			
5	Buy-and-Hold	9.73	10.31	13.04
5	Yearly Rebalancing	9.54	10.15	12.81
5	Quarterly Rebalancing	9.51	10.13	12.78
5	Monthly Rebalancing	9.53	10.13	12.82
7	Buy-and-Hold	9.96	10.37	13.11
7	Yearly Rebalancing	9.76	10.26	12.89
7	Quarterly Rebalancing	9.74	10.25	12.87
7	Monthly Rebalancing	9.76	10.26	12.91
10	Buy-and-Hold	10.24	10.73	13.37
10	Yearly Rebalancing	9.95	10.60	13.01
10	Quarterly Rebalancing	9.94	10.60	13.00
10	Monthly Rebalancing	9.96	10.61	13.04

continued

Table VI – *Continued*

Panel A: Average Annualized Semi-Volatility				
5	Buy-and-Hold	10.50	11.44	14.87
5	Yearly Rebalancing	10.21	11.17	14.39
5	Quarterly Rebalancing	10.17	11.12	14.33
5	Monthly Rebalancing	10.19	11.13	14.37
7	Buy-and-Hold	10.77	11.44	14.92
7	Yearly Rebalancing	10.48	11.23	14.44
7	Quarterly Rebalancing	10.46	11.20	14.40
7	Monthly Rebalancing	10.48	11.21	14.46
10	Buy-and-Hold	11.08	11.90	15.22
10	Yearly Rebalancing	10.68	11.71	14.58
10	Quarterly Rebalancing	10.66	11.69	14.55
10	Monthly Rebalancing	10.68	11.70	14.60

Despite the findings in Panel A of Table VI, our simulation results in Panel A of Table VII provide mixed statistical evidence that buy-and-hold exhibits a higher volatility compared with rebalancing. Only if the US markets are to be considered do we observe statistical significance at least at the 10% level. The results for the financial markets of the United Kingdom and Germany are only weakly pronounced. However, even if no significance can be detected in many cases, the positions of the 10% quantiles reported in Panel A of Table VII indicate that buy-and-hold tends to exhibit a higher average annualized volatility in all cases. An explanation is that buy-and-hold involves an increasing relative proportion of stocks, which constitute the riskier asset class compared with bonds. With an increasing time horizon, the higher volatility of stocks more and more affects the volatility of the buy-and-hold portfolio. In contrast, a periodic reallocation back to the original target weights prevents an extreme shift to riskier stocks.

2.5.2.2 Semi-Volatility

Volatility as a risk measure considers both positive and negative deviations from the sample mean. However, as positive deviations constitute an additional opportunity to generate an extra return, only the negative deviations from the mean should be economically relevant for measuring risk. Ignoring positive deviations from the mean, the semi-volatility represents a more intuitive and appropriate risk measure. Supporting our conjecture that rebalancing leads to a reduction of risk, Panel B of Table VI further reveals that buy-and-hold exhibits the highest average annualized semi-volatility for all investment horizons and all countries. Providing statistical significance for all investment horizons and all countries at least at the 10% level, Panel B of Table VII statistically substantiates the expectation that rebalancing reduces, on average, portfolio risk compared with buy-and-hold. A frequent reallocation to the predetermined portfolio weights maintains investors' initial risk-return preferences, leading to a diversification effect, and thus inducing a reduction of portfolio risk.

Table VII – CIs: Average Annualized Risk of Periodic Rebalancing

Classified by investment horizon, this table shows the confidence intervals of the difference time series between periodic rebalancing and buy-and-hold of both the average annualized volatility (Panel A) and the average annualized semi-volatility (Panel B). The sample period ranges from January 1982 to December 2011. All strategies are based on a 60% stocks and 40% bonds asset allocation with a threshold of 0%. Transaction costs are quoted at 15 bps per round-trip. BAH denotes buy-and-hold, Y yearly periodic rebalancing, Q quarterly periodic rebalancing, and M monthly periodic rebalancing. For example, M-BAH denotes the difference time series of 'Monthly periodic rebalancing minus buy-and-hold'. For each two strategies that are compared, the lower and upper boundary of the confidence interval is calculated. 1,000 simulations with an average block length of 2 are performed. Repeated simulations reveal that the results are stable. *, **, and *** represent statistical significance at the 10%, 5%, and 1% level, respectively. If there is no statistical significance, the corresponding 10% quantiles are reported.

Period	Strategy	United States		United Kingdom		Germany	
		Panel A: Average Annualized Volatility					
5	M-BAH	−0.0031	−0.0001***	−0.0029	−0.0001***	−0.0048	0.0003
5	Q-BAH	−0.0039	−0.0003***	−0.0036	−0.0002***	−0.0050	−0.0001*
5	Y-BAH	−0.0035	−0.0002***	−0.0032	−0.0002***	−0.0045	−0.0001*
7	M-BAH	−0.0035	−0.0003**	−0.0024	0.0001	−0.0050	0.0011
7	Q-BAH	−0.0040	−0.0001***	−0.0025	−0.0000*	−0.0054	0.0007
7	Y-BAH	−0.0037	−0.0001***	−0.0023	0.0000	−0.0050	0.0005
10	M-BAH	−0.0044	−0.0014*	−0.0026	0.0001	−0.0067	0.0001
10	Q-BAH	−0.0054	−0.0007***	−0.0027	0.0000	−0.0071	−0.0003*
10	Y-BAH	−0.0051	−0.0007***	−0.0025	−0.0000*	−0.0068	−0.0004*
		Panel B: Average Annualized Semi-Volatility					
5	M-BAH	−0.0054	−0.0006***	−0.0054	−0.0011***	−0.0085	−0.0013**
5	Q-BAH	−0.0056	−0.0008***	−0.0054	−0.0011***	−0.0097	−0.0005***
5	Y-BAH	−0.0051	−0.0008***	−0.0048	−0.0008***	−0.0086	−0.0004***
7	M-BAH	−0.0055	−0.0003***	−0.0042	−0.0003**	−0.0089	−0.0001**
7	Q-BAH	−0.0057	−0.0006***	−0.0042	−0.0005**	−0.0093	−0.0008**
7	Y-BAH	−0.0051	−0.0006***	−0.0039	−0.0004**	−0.0086	−0.0009**
10	M-BAH	−0.0067	−0.0009***	−0.0040	−0.0001*	−0.0112	−0.0010**
10	Q-BAH	−0.0069	−0.0011***	−0.0041	−0.0002*	−0.0116	−0.0018**
10	Y-BAH	−0.0065	−0.0012***	−0.0037	−0.0002*	−0.0121	−0.0001***

Although our results are robust in the cross-section of countries, Panel B of Table VII shows that country-specific characteristics exert a different impact on the statistical significance levels. Moreover, we observe that an increasing investment horizon leads to higher significance levels in almost all cases. Provided that statistical significance can be detected, one can infer that the calculated confidence intervals feature a higher distance to zero with increasing investment horizons. Overall, our empirical results do not confirm the findings of Arnott and Lovell (1993), who conclude from their analysis that rebalancing offers enhanced returns without increasing risk. Donohue and Yip

(2003) also report that optimal rebalancing can provide both higher returns and lower risk than other common rebalancing heuristics. Instead, we provide statistical evidence that, on average, rebalancing induces a risk reduction without sacrificing returns.

2.5.3 Risk-Adjusted Performance Measures

Both the return and the risk of a portfolio strategy are of fundamental importance to institutional investors. Therefore, we now concentrate on risk-adjusted performance measures in order to appropriately evaluate portfolio performance. In particular, our measures of interest are the Sharpe ratio, the Omega measure, and the Sortino ratio.

2.5.3.1 Sharpe Ratio

In a first step, we analyze the Sharpe ratio (Sharpe (1966)), which is the most commonly used performance measure in investment practice. Observing that there are no statistical differences between average returns, but that rebalancing leads to a significant reduction in risk, one would expect that this diversification effect also has an impact on the corresponding Sharpe ratio. Table VIII reports the average annualized Sharpe ratios of periodic rebalancing classified by strategy, investment horizon, and country. As expected, buy-and-hold leads in all cases to inferior Sharpe ratios compared with yearly, quarterly, and monthly periodic rebalancing. For example, the average Sharpe ratio of a buy-and-hold strategy using US data and assuming a 10-year investment horizon is 0.552, whereas a monthly periodic rebalancing strategy leads to an average Sharpe ratio of 0.575. With respect to quarterly and yearly periodic rebalancing, the Sharpe ratio increases to 0.579 on average.

As shown in Panel A of Table IX, these patterns are also reflected in the statistical significance levels for differences between periodic rebalancing and buy-and-hold of average annualized Sharpe ratios. In all cases, buy-and-hold is significantly outperformed by periodic rebalancing at the 1% level both for all investment horizons and for all countries under investigation. The results shown in Table VIII also illustrate the economic relevance of rebalancing for investment practice. As an example, in dependence of the country and the investment horizon under investigation, the average Sharpe ratio of a quarterly periodic rebalancing strategy increases by at least 4.28% up to 13.97% when compared to buy-and-hold. Given that rebalancing is a rule-based investment strategy, this finding is of particular importance for very large funds. Such funds are restricted by their investment decisions because the portfolio volume is too large to pursue any active management strategies. Being one of the world's largest institutional investors, Norway's Government Pension Fund Global provides a good example of conducting rebalancing as a rule-based investment strategy.

Table VIII – Average Annualized Sharpe Ratios

Classified by investment horizon, this shows the average annualized Sharpe ratios of buy-and-hold as well as periodic rebalancing with yearly, quarterly, and monthly trading intervals over the sample period from January 1982 to December 2011. All strategies are based on a 60% stocks and 40% bonds asset allocation with a threshold of 0%. Transaction costs are quoted at 15 bps per round-trip. 1,000 simulations with an average block length of 2 are performed. Repeated simulations reveal that the results are stable.

Period	Rebalancing Strategy	United States	United Kingdom	Germany
5	Buy-and-Hold	0.554	0.336	0.315
5	Yearly Rebalancing	0.580	0.354	0.354
5	Quarterly Rebalancing	0.583	0.356	0.359
5	Monthly Rebalancing	0.580	0.355	0.356
7	Buy-and-Hold	0.571	0.379	0.330
7	Yearly Rebalancing	0.597	0.398	0.369
7	Quarterly Rebalancing	0.598	0.399	0.372
7	Monthly Rebalancing	0.594	0.398	0.368
10	Buy-and-Hold	0.552	0.374	0.314
10	Yearly Rebalancing	0.579	0.389	0.355
10	Quarterly Rebalancing	0.579	0.390	0.356
10	Monthly Rebalancing	0.575	0.389	0.351

In order to substantiate our hypothesis that rebalancing generates a value added for institutional investors, we additionally test threshold and range rebalancing strategies by implementing a symmetric no-trade region around the target weights. Once a rebalancing threshold is introduced, there are two cases that need to be distinguished with regard to the practical implementation. For the first alternative strategy, a strict adjustment to the target weights (Buetow et al. (2002), Harjoto and Jones (2006)) is required when an asset exceeds the predetermined interval boundaries within a given interval. Threshold rebalancing is captured by strategies (5)-(7) in Table II. In contrast, the second alternative rebalancing strategy requires a rebalancing back to the nearest edge of the given threshold rather than to the initial portfolio weights (Leland (1999)). Range rebalancing refers to strategies (8)-(10) in Table II. Referring to Norway's Government Pension Fund Global, we assume a symmetric no-trade interval of ±3% (Norwegian Ministry of Finance (2012)). Panel B and Panel C of Table IX show the confidence intervals for these two alternative rebalancing strategies. Confirming our previous results for the simpler periodic rebalancing strategy, a buy-and-hold strategy is significantly dominated by both threshold rebalancing and range rebalancing in terms of average Sharpe ratios at all rebalancing frequencies. Without any exception, the difference is always significant at the 1% level. In results not shown, this finding is even robust when the threshold is changed to ±10%. Accordingly, we conclude from our investigation of the Sharpe ratio that the dominance of rebalancing over buy-and-hold is independent of the choice of a specific rebalancing strategy.

Table IX – CIs: Average Annualized Sharpe Ratio

Classified by investment horizon, this table shows the confidence intervals of the difference time series of the average Sharpe ratio of periodic rebalancing (Panel A), threshold rebalancing (Panel B), and range rebalancing (Panel C). The sample period ranges from January 1982 to December 2011. All strategies are based on a 60% stocks and 40% bonds asset allocation with a threshold of 3% (if applicable). Transaction costs are quoted at 15 bps per round-trip. BAH denotes buy-and-hold, Y yearly rebalancing, Q quarterly rebalancing, and M monthly rebalancing. For example, M-BAH denotes the difference time series of 'Monthly rebalancing minus buy-and-hold'. For each two strategies that are compared, the lower and upper boundary of the confidence interval is calculated. 1,000 simulations with an average block length of 2 are performed. Repeated simulations reveal that the results are stable. *, **, and *** represent statistical significance at the 10%, 5%, and 1% level, respectively. If there is no statistical significance, the corresponding 10% quantiles are reported.

Period	Strategy	United States		United Kingdom		Germany	
Panel A: Average Annualized Sharpe Ratio of Periodic Rebalancing							
5	M-BAH	0.0151	0.0366***	0.0111	0.0263***	0.0301	0.0522***
5	Q-BAH	0.0181	0.0390***	0.0122	0.0276***	0.0332	0.0552***
5	Y-BAH	0.0139	0.0367***	0.0101	0.0253***	0.0268	0.0518***
7	M-BAH	0.0112	0.0347***	0.0090	0.0272***	0.0243	0.0512***
7	Q-BAH	0.0153	0.0387***	0.0099	0.0289***	0.0285	0.0553***
7	Y-BAH	0.0145	0.0382***	0.0086	0.0274***	0.0267	0.0533***
10	M-BAH	0.0072	0.0368***	0.0020	0.0246***	0.0216	0.0513***
10	Q-BAH	0.0121	0.0405***	0.0038	0.0259***	0.0267	0.0562***
10	Y-BAH	0.0132	0.0416***	0.0041	0.0252***	0.0252	0.0565***
Panel B: Average Annualized Sharpe Ratio of Threshold Rebalancing							
5	M-BAH	0.0156	0.0366***	0.0115	0.0261***	0.0297	0.0522***
5	Q-BAH	0.0162	0.0382***	0.0115	0.0268***	0.0315	0.0544***
5	Y-BAH	0.0131	0.0356***	0.0093	0.0245***	0.0256	0.0511***
7	M-BAH	0.0129	0.0364***	0.0083	0.0276***	0.0255	0.0515***
7	Q-BAH	0.0143	0.0378***	0.0090	0.0283***	0.0278	0.0542***
7	Y-BAH	0.0136	0.0369***	0.0076	0.0269***	0.0257	0.0524***
10	M-BAH	0.0089	0.0380***	0.0023	0.0245***	0.0229	0.0519***
10	Q-BAH	0.0131	0.0398***	0.0034	0.0255***	0.0267	0.0557***
10	Y-BAH	0.0120	0.0413***	0.0046	0.0246***	0.0250	0.0558***
Panel C: Average Annualized Sharpe Ratio of Threshold Rebalancing							
5	M-BAH	0.0128	0.0334***	0.0093	0.0233***	0.0284	0.0498***
5	Q-BAH	0.0124	0.0323***	0.0089	0.0225***	0.0267	0.0484***
5	Y-BAH	0.0089	0.0260***	0.0060	0.0180***	0.0186	0.0412***
7	M-BAH	0.0114	0.0344***	0.0074	0.0250***	0.0249	0.0507***
7	Q-BAH	0.0118	0.0340***	0.0070	0.0244***	0.0239	0.0501***
7	Y-BAH	0.0091	0.0294***	0.0052	0.0209***	0.0195	0.0441***
10	M-BAH	0.0102	0.0380***	0.0027	0.0233***	0.0245	0.0531***
10	Q-BAH	0.0113	0.0377***	0.0035	0.0231***	0.0244	0.0531***
10	Y-BAH	0.0096	0.0351***	0.0029	0.0207***	0.0206	0.0492***

As Table I shows, the stock markets of the United States, the United Kingdom, and Germany all exhibit a negative skewness, implying a higher probability of large losses compared with symmetric return distributions. Furthermore, in results not shown, buy-

Table X – Skewness and Kurtosis of Rebalancing Strategies

Classified by rebalancing strategy and country under investigation, this table presents the skewness and the kurtosis over the entire sample period from January 1982 to December 2011. All strategies are based on a 60% stocks and 40% bonds asset allocation with a threshold of 0%. Transaction costs are quoted at 15 bps per round-trip. The skewness and kurtosis are calculated as the third and fourth normalized centered moments.

Panel A: Skewness			
Rebalancing Strategy	United States	United Kingdom	Germany
Buy-and-Hold	-0.75	-0.86	-0.87
Yearly Rebalancing	-0.62	-0.68	-0.56
Quarterly Rebalancing	-0.50	-0.51	-0.59
Monthly Rebalancing	-0.50	-0.52	-0.66
Panel B: Kurtosis			
Rebalancing Strategy	United States	United Kingdom	Germany
Buy-and-Hold	5.09	6.44	5.19
Yearly Rebalancing	5.52	5.96	4.44
Quarterly Rebalancing	4.82	4.71	4.55
Monthly Rebalancing	4.80	4.69	4.81

and-hold exhibits a higher average stock allocation compared with rebalancing. This finding is again robust for all countries, all investment horizons, and all rebalancing strategies. Accordingly, one would expect that buy-and-hold resembles this higher negative skewness due to the higher relative proportion of the riskier asset class stocks. For the same reason, we also hypothesize a higher kurtosis for buy-and-hold as extreme returns should be much more pronounced. Applying buy-and-hold as well as periodic rebalancing to the entire 30-year sample period, Table X confirms these expectations concerning the skewness. With the exception of yearly periodic rebalancing, we observe a similar pattern for the kurtosis. Despite the widespread use in investment practice, the Sharpe ratio suffers from the shortcoming of assuming investors that are characterized by mean-variance preferences.[14] Given the distributional characteristics of the different strategies, we thus consider the Omega measure and the Sortino ratio as downside risk-adjusted performance measures.

2.5.3.2 Omega Measure

The Omega measure (Shadwick and Keating (2002)) represents a special case of the more general performance measure Kappa (Kaplan and Knowles (2004)). It is characterized by the ratio of gains to losses relative to a predefined target return required by

[14] Adcock et al. (2012) show that even if returns are not normally distributed, the rank correlation of the Sharpe ratio and other performance measures is one or very close to it.

the investor:

$$\Omega_i(\tau) = \frac{\int_{\tau}^{\infty}(1 - F(r_i))dr_i}{\int_{-\infty}^{\tau} F(r_i)dr_i}, \tag{7}$$

where τ denotes the investor's required rate of return and $F(r_i)$ is the cumulative distribution function of the monthly return r of strategy i. For simplicity, we set the target return to zero, which allows us to distinguish between realized gains and losses. In contrast to the Sharpe ratio, the Omega measure makes no assumptions about the shape of the distribution. Furthermore, by considering the entire return distribution, all moments are taken into account. Panel A of Table XI reconfirms the empirical results in Panel A of Table IX (Sharpe ratio). Buy-and-hold is outperformed by all periodic rebalancing strategies in terms of average Omega measures at the 1% level. These findings apply to all analyzed investment horizons and all countries.

2.5.3.3 Sortino Ratio

Similar to the Omega measure, the Sortino ratio (Sortino and Price (1994)) is also a special case of Kappa. In contrast to the Sharpe ratio, it penalizes only those returns that fall below a target return:

$$S_i(\tau) = \frac{\bar{r}_i - \tau}{\sqrt{\int_{-\infty}^{\tau}(\tau - r_i)^2 f(r_i)dr_i}}, \tag{8}$$

where $\bar{r}_i$ denotes the average return of the underlying strategy i, $f(r_i)$ the corresponding probability density function, and τ the required target return of the investor.

As investors should be concerned about negative, but not about positive deviations relative to a required rate of return, the Sortino ratio may be a more appropriate concept to measure risk-adjusted returns than the common Sharpe ratio. Setting the target return again to zero, we obtain qualitatively very similar results compared with Panel A of Table IX (Sharpe ratio) and Panel A of Table XI (Omega measure). Therefore, the empirical results of the Sortino ratios in Panel B of Table XI strongly reconfirm our hypothesis that rebalancing generates a value added for institutional investors. Moreover, when testing both the Omega measure and the Sortino ratio for threshold and range rebalancing with a no-trade region of ±3% against buy-and-hold, we again observe statistical significance at the 1% level (not tabulated) for all investment horizons and all countries.

Table XI – CIs: Average Downside Risk-Adjusted Performance of Periodic Rebalancing
Classified by investment horizon, this table shows the confidence intervals of the difference time series between periodic rebalancing and buy-and hold of both the average Omega measure (Panel A) and the average annualized Sortino ratio (Panel B). The sample period ranges from January 1982 to December 2011. All strategies are based on a 60% stocks and 40% bonds asset allocation with a threshold of 0%. Transaction costs are quoted at 15 bps per round-trip. BAH denotes buy-and-hold, Y yearly periodic rebalancing, Q quarterly periodic rebalancing, and M monthly periodic rebalancing. For example, M-BAH denotes the difference time series of 'Monthly periodic rebalancing minus buy-and-hold'. For each two strategies that are compared, the lower and upper boundary of the confidence interval is calculated. 1,000 simulations with an average block length of 2 are performed. Repeated simulations reveal that the results are stable. *, **, and *** represent statistical significance at the 10%, 5%, and 1% level, respectively. If there is no statistical significance, the corresponding 10% quantiles are reported.

Period	Strategy	United States		United Kingdom		Germany	
			Panel A: Average Omega Measure				
5	M-BAH	0.0547	0.1383***	0.0436	0.1145***	0.0509	0.1203***
5	Q-BAH	0.0580	0.1383***	0.0451	0.1152***	0.0549	0.1192***
5	Y-BAH	0.0468	0.1169***	0.0355	0.0979***	0.0438	0.1023***
7	M-BAH	0.0387	0.1141***	0.0254	0.1009***	0.0404	0.1110***
7	Q-BAH	0.0444	0.1186***	0.0290	0.1025***	0.0453	0.1154***
7	Y-BAH	0.0407	0.1134***	0.0260	0.0943***	0.0414	0.1055***
10	M-BAH	0.0326	0.1095***	0.0081	0.0816***	0.0313	0.1038***
10	Q-BAH	0.0393	0.1150***	0.0109	0.0854***	0.0377	0.1096***
10	Y-BAH	0.0373	0.1148***	0.0105	0.0794***	0.0360	0.1041***
			Panel B: Average Annualized Sortino Ratio				
5	M-BAH	0.0797	0.1873***	0.0623	0.1511***	0.0769	0.1726***
5	Q-BAH	0.0843	0.1898***	0.0652	0.1524***	0.0825	0.1741***
5	Y-BAH	0.0684	0.1619***	0.0511	0.1301***	0.0681	0.1518***
7	M-BAH	0.0661	0.1667***	0.0420	0.1443***	0.0613	0.1618***
7	Q-BAH	0.0737	0.1737***	0.0468	0.1454***	0.0703	0.1675***
7	Y-BAH	0.0664	0.1649***	0.0408	0.1324***	0.0620	0.1566***
10	M-BAH	0.0553	0.1698***	0.0221	0.1231***	0.0512	0.1541***
10	Q-BAH	0.0659	0.1759***	0.0258	0.1265***	0.0606	0.1654***
10	Y-BAH	0.0610	0.1768***	0.0233	0.1188***	0.0577	0.1576***

Overall, our simulation set-up allows us to determine whether a rebalancing strategy is able to generate a value added in comparison to buy-and-hold and to identify the source of this value added. Although there are no statistical differences in average returns between buy-and-hold and rebalancing, we provide statistical evidence that the reduction of risk is the major source of the value added of rebalancing. Considering both the return and the risk of a given strategy, we further document that the Sharpe ratio, the Omega measure, and the Sortino ratio – as simple measures of value added – of all rebalancing strategies are, on average, significantly higher compared with buy-and-hold. In conclusion, while the return effect is not responsible for the superiority of the applied risk-adjusted performance measures, it is a risk management argument that justifies the widespread use of rebalancing strategies in the asset management practice.

2.6 Robustness Checks

2.6.1 Transaction Costs

When evaluating portfolio performance, transaction costs represent a crucial factor that must not be ignored. Indeed, it is of utmost importance for institutional portfolio management that realistic transaction costs are considered in order to draw meaningful economic conclusions. Otherwise, a statistical comparison between rebalancing and buy-and-hold may lead to biases which result in suboptimal portfolio recommendations. Although we assume realistic transaction costs of 15 bps per round-trip throughout our entire analysis (see footnote 7 for a discussion), we also apply unrealistic high transaction costs of 100 bps in order to show their impact on rebalancing. Even though these higher transaction costs negatively influence portfolio returns of all rebalancing strategies, the main empirical findings remain qualitatively almost unchanged (results not shown). In all but two cases, all rebalancing strategies consistently outperform buy-and-hold at the 1% level in terms of average risk-adjusted portfolio performance. With respect to monthly periodic rebalancing, we only observe statistical significance at the 5% level for the U.S. market and no significance for the U.K. market.

2.6.2 Variation of the No-Trade Interval for Threshold and Range Rebalancing

Analyzing the value added of rebalancing, a resulting question of interest is how large the symmetric no-trade region must be in order to loose statistical significance. The larger the no-trade interval is, the higher the similarity between both threshold and range rebalancing compared with buy-and-hold will be. Given a symmetric threshold of ±10% and an investment horizon of 10 years, all threshold and range rebalancing strategies are highly significant for all countries (not tabulated). In contrast, examining a symmetric threshold of ±30%, there is no statistical significance at all. Investigating the German financial market, we observe statistical significance regarding a threshold of ±20%, whereas no significance of either rebalancing class can be detected for the U.K. financial market concerning a threshold of about ±15% (for both threshold and range rebalancing). Analyzing the financial market of the United States, statistical significance highly depends on both the rebalancing class and the trading frequency and can be detected for threshold levels between ±15% and ±20%.

2.6.3 Variation of the Average Block Length

According to Ledoit and Wolf (2008), the stationary bootstrap of Politis and Romano (1994) is quite insensitive to the choice of the average block length. Nevertheless, we also apply average block lengths of $b \in \{4,6,8\}$ as a final robustness check. For example, analyzing an investment horizon of 10 years, Ledoit and Wolf (2008, 2011) estimate an optimal block size of $b = 4$ for their mutual funds application and $b = 6$ for their hedge fund application.[15] Therefore, our selection of the average block size of $b = 2$ as well as of $b \in \{4,6,8\}$ should be statistically and economically reasonable. In particular, longer block lengths lead to confidence intervals with a higher tendency of including 0, which makes it even more difficult to find evidence for statistical significance. In conclusion, our primary result that rebalancing outperforms buy-and-hold in terms of average risk-adjusted performance remains qualitatively unchanged (not reported).

2.7 Conclusion

This study makes two important contributions to the academic literature as well as to investment practice by examining the issue of why institutional investors prefer rebalancing over buy-and-hold, even though rebalancing strategies require the selling of a fraction of the better performing assets and investing the proceeds in the weaker performing assets. First of all, we provide strong empirical evidence that all rebalancing strategies (periodic, threshold, and range rebalancing with yearly, quarterly, and monthly trading intervals) significantly outperform buy-and-hold for all countries (United States, United Kingdom, and Germany) and for all investment horizons (5, 7, and 10 years), regardless of whether the Sharpe ratio, the Omega measure, or the Sortino ratio is considered for evaluating risk-adjusted portfolio performance. Accordingly, we infer from our analysis that the dominance of rebalancing over buy-and-hold is independent of the choice of a specific rebalancing strategy. Moreover, our results are also of meaningful economic importance. For example, in comparison to buy-and-hold, a quarterly periodic rebalancing strategy leads to an increase of the average Sharpe ratio of about 5% for the financial markets of the United States and the United Kingdom and almost 13% for the German financial market. Secondly, we document that while there are no statistical differences in average returns between rebalancing and buy-and-hold, the superior risk-adjusted performance of rebalancing is attributable to reduced risk. The regular reallocation to the original asset allocation prevents an extreme drift from the weaker performing but less risky asset class towards the better performing but

[15] However, using a semi-parametric model to fit the observed return data, the algorithm for identifying the optimal block size is different from the method suggested by Politis and White (2004) as well as by Patton et al. (2009).

more risky one. As a result, the portfolio risk is not only significantly reduced, but the original portfolio diversification is also almost entirely preserved. Overall, we conclude that it is a risk management argument which justifies the widespread use of rebalancing in investment practice: The primary objective of any rebalancing strategy is the reduction of risk with respect to a given asset allocation.

References

Adcock, C.J., N. Areal, M.J.R. Armada, M.C. Cortez, B. Oliveira, and F. Silva (2012). "Conditions under Which Portfolio Performance Measures Are Monotonic Functions of the Sharpe Ratio." Working Paper, University of Sheffield Management School, Sheffield, 2012.

Annaert, J., S. Van Osselaer, and B. Verstraete (2009). "Performance Evaluation of Portfolio Insurance Strategies Using Stochastic Dominance Criteria." In: *Journal of Banking and Finance* 33, pp. 272–280.

Arnott, R.D. and R.M. Lovell (1993). "Rebalancing: Why? When? How Often?" In: *Journal of Investing* 2, pp. 5–10.

Avramov, D. and R. Wermers (2006). "Investing in Mutual Funds When Returns Are Predictable." In: *Journal of Financial Economics* 81, pp. 339–377.

Ball, C. and W.N. Torous (2000). "Stochastic Correlation Across International Stock Markets." In: *Journal of Empirical Finance* 7, pp. 373–388.

Barras, L., O. Scaillet, and R. Wermers (2010). "False Discoveries in Mutual Fund Performance: Measuring Luck in Estimated Alphas." In: *Journal of Finance* 65, pp. 179–216.

Brennan, M., F. Li, and W.N. Torous (2005). "Dollar Cost Averaging." In: *Review of Finance* 9, pp. 509–535.

Brock, W., J. Lakonishok, and B. LeBaron (1992). "Simple Technical Trading Rules and the Stochastic Properties of Stock Returns." In: *Journal of Finance* 47, pp. 1731–1764.

Buetow, G.W., E. Hunt, R. Sellers, D. Trotter, and W.A. Whipple (2002). "The Benefits of Rebalancing." In: *Journal of Portfolio Management* 28, pp. 23–32.

Buraschi, A., P. Porchia, and F. Trojani (2010). "Correlation Risk and Optimal Portfolio Choice." In: *Journal of Finance* 65, pp. 393–420.

Busse, J. A., A. Goyal, and S. Wahal (2010). "Performance and Persistence in Institutional Investment Management." In: *Journal of Finance* 65, pp. 765–790.

Campbell, J.Y. and L.M. Viceira (2002). *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors*. Oxford University Press.

Carhart, M.M. (1997). "On Persistence in Mutual Fund Performance." In: *Journal of Finance* 52, pp. 57–82.

Chambers, D., E. Dimson, and A. Ilmanen (2012). "The Norway Model." In: *Journal of Portfolio Management* 38, pp. 67–81.

Clark, P.K. (1973). "A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices." In: *Econometrica* 41, pp. 135–155.

Cremers, K.J.M. and A. Petajisto (2009). "How Active Is Your Fund Manager? A New Measure That Predicts Performance." In: *Review of Financial Studies* 22, pp. 3329–3365.

De Bondt, W.F.M. and R. Thaler (1985). "Does the Stock Market Overreact?" In: *Journal of Finance* 40, pp. 793–805.

Donohue, C. and K. Yip (2003). "Optimal Portfolio Rebalancing with Transaction Costs." In: *Journal of Portfolio Management* 29, pp. 49–63.

Dumas, B. and B. Solnik (1995). "The World Price of Foreign Exchange Risk." In: *Journal of Finance* 50, pp. 445–479.

Efron, B. and R.J. Tibshirani (1998). *An Introduction to the Bootstrap*. Chapman & Hall/CRC.

Engle, R.F. (1982). "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation." In: *Econometrica* 50, pp. 987–1007.

Engle, R.F., D.M. Lilien, and R.P. Robins (1987). "Estimating Time-Varying Risk Premia in the Term Structure: The ARCH-M Model." In: *Econometrica* 55, pp. 391–408.

Eraker, B. (2004). "Do Stock Prices and Volatility Jump? Reconciling Evidence from Spot and Option Prices." In: *Journal of Finance* 59, pp. 1367–1403.

Erb, C.B., C.R. Harvey, and T.E. Viskanta (1994). "Forecasting International Equity Correlations." In: *Financial Analysts Journal* 50, pp. 32–45.

Fama, E.F. (1965). "The Behavior of Stock Market Prices." In: *Journal of Business* 38, pp. 34–105.

Fama, E.F. and K.R. French (2010). “Luck versus Skill in the Cross-Section of Mutual Fund Returns.” In: *Journal of Finance* 65, pp. 1915–1947.

Ferson, W.E. and C.R. Harvey (1991). “The Variation of Economic Risk Premiums.” In: *Journal of Political Economy* 99, pp. 385–415.

– (1993). “The Risk and Predictability of International Equity Returns.” In: *Review of Financial Studies* 6, pp. 527–566.

Ferson, W.E., S. Kandel, and R.F. Stambaugh (1987). “Tests of Asset Pricing with Time-Varying Expected Risk Premiums and Market Betas.” In: *Journal of Finance* 42, pp. 201–220.

French, K.R. (2008). “Presidential Address: The Cost of Active Investing.” In: *Journal of Finance* 63, pp. 1537–1573.

Gibbons, M.R. and W.E. Ferson (1985). “Testing Asset Pricing Models with Changing Expectations and an Unobservable Market Portfolio.” In: *Journal of Financial Economics* 14, pp. 217–236.

Harjoto, M.A. and F.J. Jones (2006). “Rebalancing Strategies for Stocks and Bonds Asset Allocation.” In: *Journal of Wealth Management* 9, pp. 37–44.

Hendricks, D., J. Patel, and R. Zeckhauser (1993). “Hot Hands in Mutual Funds: Short-Run Persistence of Performance, 1974-1988.” In: *Journal of Finance* 48, pp. 93–130.

Ingersoll, J., M. Spiegel, W. Goetzmann, and I. Welch (2007). “Portfolio Performance Manipulation and Manipulation-Proof Performance Measures.” In: *Review of Financial Studies* 20, pp. 1503–1546.

Jaconetti, C.M., F.M. Kinniry, and Y. Zilbering (2010). *Best Practices for Portfolio Rebalancing*. Tech. rep. Vanguard.

Jones, S.K. and J.B. Stine (2010). “Expected Utility and the Non-Normal Returns of Common Portfolio Rebalancing Strategies.” In: *Journal of Asset Management* 10, pp. 406–419.

Kaplan, P.D. and J.A. Knowles (2004). “Kappa: A Generalized Downside Risk-Adjusted Performance Measure.” In: *Journal of Performance Measurement* 8, pp. 42–54.

Kimball, M.S., M.D. Shapiro, T. Shumway, and J. Zhang (2011). “Portfolio Rebalancing in General Equilibrium.” Working Paper, University of Michigan, Michigan, 2011.

Kosowski, R., A. Timmermann, R. Wermers, and H. White (2006). "Can Mutual Fund "Stars" Really Pick Stocks? New Evidence from a Bootstrap Analysis." In: *Journal of Finance* 61, pp. 2551–2595.

Ledoit, O. and M. Wolf (2008). "Robust Performance Hypothesis Testing with the Sharpe Ratio." In: *Journal of Empirical Finance* 15, pp. 850–859.

– (2011). "Robust Performance Hypothesis Testing with Variance." In: *Wilmott Magazine* 55, pp. 86–89.

Leland, H. E. (1999). "Optimal Portfolio Management with Transactions Costs and Capital Gains Taxes." Research Program in Finance Working Paper RPF-290, University of California, Berkeley, 1999.

Longin, F. and B. Solnik (2001). "Extreme Correlation of International Equity Markets." In: *Journal of Finance* 56, pp. 649–676.

Mandelbrot, B.B. (1963). "The Variation of Certain Speculative Prices." In: *Journal of Business* 36, pp. 394–419.

Masters, S.J. (2003). "Rebalancing." In: *Journal of Portfolio Management* 29, pp. 52–57.

McLellan, R., W. Kinlaw, and E. Abouzaid (2009). "Equity Turbulence, Fixed Income Illiquidity, and Portfolio Reallocation: The Case for Synthetic Rebalancing." In: *Transition Management Guide* 2, pp. 18–27.

Merton, R. (1971). "Optimum Consumption and Portfolio Rules in a Continuous-Time Model." In: *Journal of Economic Theory* 3, pp. 373–413.

National Bureau of Economic Research (2012). *US Business Cycle Expansions and Contractions*. URL: http://www.nber.org/cycles/cyclesmain.html..

Ng, L. (1991). "Tests of the CAPM with Time-Varying Covariances: A Multivariate GARCH Approach." In: *Journal of Finance* 46, pp. 1507–1521.

Norwegian Ministry of Finance (2012). *Report No. 17 to the Storting: The Management of the Government Pension Fund in 2011*. URL: http://www.regjeringen.no/pages/37635499/stm_gpf.pdf.

Patton, A., D.N. Politis, and H. White (2009). "Correction to "Automatic Block-Length Selection for the Dependent Bootstrap" by D. Politis and H. White." In: *Econometric Reviews* 28, pp. 372–375.

Perold, A.F. and W.F. Sharpe (1988). "Dynamic Strategies for Asset Allocation." In: *Financial Analysts Journal* 44, pp. 16–27.

Plaxco, L.M. and R.D. Arnott (2002). "Rebalancing a Global Policy Benchmark." In: *Journal of Portfolio Management* 28, pp. 9–22.

Politis, D.N. and J.P. Romano (1994). "The Stationary Bootstrap." In: *Journal of the American Statistical Association* 89, pp. 1303–1313.

Politis, D.N. and H. White (2004). "Automatic Block-Length Selection for the Dependent Bootstrap." In: *Econometric Reviews* 23, pp. 53–70.

Poterba, J.M. and L.H. Summers (1988). "Mean Reversion in Stock Prices: Evidence and Implications." In: *Journal of Financial Economics* 22, pp. 27–59.

Romano, J.P. and M. Wolf (2006). "Improved Nonparametric Confidence Intervals in Time Series Regressions." In: *Journal of Nonparametric Statistics* 18, pp. 199–214.

Shadwick, W.F. and C. Keating (2002). "A Universal Performance Measure." In: *Journal of Performance Measurement* 6, pp. 59–84.

Sharpe, W.F. (1966). "Mutual Fund Performance." In: *Journal of Business* 39, pp. 119–138.

– (1991). "The Arithmetic of Active Management." In: *Financial Analysts Journal* 47, pp. 7–9.

– (2010). "Adaptive Asset Allocation Policies." In: *Financial Analysts Journal* 66, pp. 45–59.

Sortino, F.A. and L.N. Price (1994). "Performance Measurement in a Downside Risk Framework." In: *Journal of Investing* 3, pp. 59–64.

Sun, W., A.C. Fan, L.-W. Chen, T. Schouwenaars, and M.A. Albota (2006). "Optimal Rebalancing for Institutional Portfolios." In: *Journal of Portfolio Management* 32, pp. 33–43.

Tokat, Y. and N. Wicas (2007). "Portfolio Rebalancing in Theory and Practice." In: *Journal of Investing* 16, pp. 52–59.

Wermers, R. (2000). "Mutual Fund Performance: An Empirical Decomposition into Stock-Picking Talent, Style, Transaction Costs, and Expenses." In: *Journal of Finance* 55, pp. 1655–1695.

Chapter 3

Testing Rebalancing Strategies for Stock-Bond Portfolios: What Is the Optimal Rebalancing Strategy?

with H. Dichtl and W. Drobetz

Acknowledgements

We appreciate helpful discussions with participants at the 2013 Midwest Finance Association Annual Meeting in Chicago.

Abstract

We compare the performance of different rebalancing strategies under realistic market conditions by reporting statistical significance levels. Our analysis is based on historical data from the United States, the United Kingdom, as well as Germany and comprises three different classes of rebalancing (namely *periodic*, *threshold*, and *range rebalancing*). Despite cross-country differences, we provide evidence that both excessive as well as too infrequent rebalancing lead to an inferior risk-adjusted portfolio performance. Specifically, the optimal rebalancing strategy seems to be quarterly periodic rebalancing for all three countries under investigation.

3.1 Introduction

Once an investor's risk and return preferences as well as his regulatory environment have been identified, it is the primary objective of any institutional asset manager to implement and supervise the most suitable asset allocation for his client. Given this initial asset allocation, the literature differentiates between three reasons for portfolio rebalancing: (i) rebalancing due to a shift in an investor's risk profile and/or modified regulatory requirements; (ii) rebalancing based on changes in the expectations of future returns and risks; and (iii) rebalancing due to market movements. As discussed in Fabozzi et al. (2006) as well as in Leibowitz and Bova (2011), the first two reasons require the asset manager to construct a new optimal portfolio.

In this study, we focus on the third reason, which legitimates portfolio rebalancing as a cost-efficient, rule-based investment strategy. As different assets generate different rates of return, a portfolio's relative asset composition will deviate from the target weights over time. In order to remain consistent with the institutional investor's initially evaluated risk and return preferences, the portfolio manager needs to rebalance the assets back to their predefined target weights. The resulting research question is which rebalancing algorithm and which rebalancing frequency should be adopted. In short, what is the optimal rebalancing strategy?

This issue is of considerable importance for investment practice as exemplarily documented by the Norwegian Government Pension Fund Global (GPFG). Being one of the world's largest institutional investors with 554.96 billion US$ AuM by the end of December 2011, the GPFG is a good example of pursuing rebalancing as a cost-efficient, rule-based investment strategy (Norwegian Ministry of Finance (2012)). In search of the optimal risk-return reward, even small deviations from the strategic asset allocation could be economically relevant. In particular, this applies to very large funds which are restricted by their investment decisions according to their size. With 19.34 billion US$ AuM by the end of June 2012, the Yale endowment also conducts rebalancing with the primary goal of maintaining the original risk profile as well as exploiting return-generating opportunities caused by excess security price volatility (Swensen (2009), The Yale Endowment (2012)).

Examining dynamic portfolio strategies that invest only in stocks and bills, Perold and Sharpe (1988) have laid the theoretical foundations for our empirical analysis. They point out that the upside potential of a buy-and-hold strategy is proportional to the amount allocated into stocks, while its downside protection is proportional to the amount allocated into bills. Analyzing rebalancing strategies, three important conclusions can be drawn from the regular reallocation to the weaker performing asset class. First of all, rebalancing exhibits a lower upside potential in comparison to buy-and-hold during a persistent market upswing. Secondly, representing the sale of portfolio insurance, rebalancing also provides less downside protection in persistent market downswings. Nevertheless, rebalancing performs best in volatile markets that feature neither a persistent market downswing nor a persistent market upswing. According to Perold and Sharpe (1988), these market conditions advantage rebalancing strategies, which may ultimately result in both improved portfolio returns and a reduction of portfolio risk.

Investigating the average return, the volatility, and the Treynor ratio of several rebalancing strategies over the period from 1968 to 1991, Arnott and Lovell (1993) document that a monthly rebalancing strategy features the highest return while the corresponding volatility is only slightly higher compared to the strategy with the lowest volatility. However, using the Treynor ratio as a performance measure that incorporates both a strategy's return and its systematic risk, the empirical results are weaker. In fact, during the underlying 24-year sample period, all Treynor ratios lie very close together within the interval (0.784; 0.794), and thus it is not obvious which strategy actually performs best. Nevertheless, inferring from their analysis that rebalancing offers enhanced returns without increasing risk, Arnott and Lovell (1993) recommend a monthly rebalancing strategy to investors with a long investment horizon.

Evaluating the performance on the basis of the Sharpe ratio over the period from 1986 to 2000 for different risk-profiles, Tsai (2001) shows that a frequent reallocation back to the target weights seems to provide some value added to institutional investors. However, as no single strategy is consistently better across portfolios of different risk profiles, Tsai (2001) argues that it does not matter much which rebalancing strategy is adopted.

Examining the period from 1995 to 2004, Harjoto and Jones (2006) report that a rebalancing strategy with an incorporated no-trade interval of 15% leads to both the highest average return and the lowest standard deviation, which in turn also results in the highest Sharpe ratio. This empirical finding also remains valid when the sample period is divided into an economic boom, a bust, and a recovery subsample. Taken as a whole, Harjoto and Jones (2006) conclude that investors should readjust their portfolio structure, though not too frequently. Nevertheless, three potential drawbacks are worth noting: (i) The analysis is based on one single 10-year period, which intensifies the potential problem of data snooping; (ii) transaction costs should have been incorporated because they might have a major influence on any reallocation decisions; (iii) the bust and recovery periods may not represent suitable estimators as they are based on only 27 and 30 observations, respectively.

Analyzing rebalancing strategies over the period from 1926 to 2009, Jaconetti et al. (2010) show that buy-and-hold exhibits the highest average annualized return with a value of 9.1% after an investment period of 84 years, but also the highest volatility with a value of 14.4% due to an average stock allocation of 84.1%. All remaining rebalancing strategies feature average returns that differ slightly, ranging between 8.5% and 8.8%. The standard deviations also lie within a narrow band of 11.8% to 12.3%. While it is evident that most institutional investors cannot apply a buy-and-hold strategy on a long-term basis, it is not obvious which rebalancing strategy leads to superior results. Accordingly, Jaconetti et al. (2010) conclude that there is no universally optimal rebalancing strategy.

The mixed results of the studies presented above can be explained by the path dependency of rebalancing, which affects all dynamic portfolio strategies. As capital markets do not exhibit arbitrage opportunities over prolonged periods of time, this path dependency further implies that there is no particular rebalancing strategy that features a better risk-return reward in all market environments in comparison to any other rebalancing strategy. However, the question that remains to be answered is whether a specific rebalancing strategy leads to a higher risk-adjusted performance on average.

Our first contribution to the literature relates to the implemented methodological approach, which is based on the stationary bootstrap of Politis and Romano (1994) and enables us to compare different rebalancing strategies with each other on a statistical basis by reporting statistical significance levels. Secondly, academic literature has so far remained incomplete by having excluded analyses of rebalancing strategies with a focus on institutional investors outside the US. For this reason, we do not only examine the financial markets of the United States, but also those of the United Kingdom and Germany in order to check whether country-specific characteristics have an impact on the performance of rebalancing.

Our findings have immediate practical implications. First of all, evaluating risk-adjusted portfolio performance on the basis of the Sharpe ratio, the Sortino ratio, and the Omega measure, we provide evidence that both excessive rebalancing (monthly periodic rebalancing) as well as too infrequent rebalancing (yearly range rebalancing) provoke inferior results, thus pointing out that there may be an optimal rebalancing strategy. Secondly, the optimal trading patterns change with respect to the underlying rebalancing algorithm. Within the corresponding rebalancing class, quarterly periodic, quarterly threshold, and monthly range rebalancing seem to produce the highest risk-adjusted performance. Thirdly, quarterly periodic rebalancing significantly outperforms quarterly threshold rebalancing as well as monthly range rebalancing for all countries and all investment horizons. Overall, our history-based simulation results provide strong evidence that quarterly periodic rebalancing tends to be the optimal rebalancing strategy for all three analyzed countries. We further substantiate this finding by analyzing 10-year investment horizons that all have been realized in the past with the help of a rolling window. Fourthly, short-term momentum seems to be the primary source capable of explaining the statistically significant differences between monthly and quarterly periodic rebalancing.

The remainder of this study is structured as follows: Section 3.2 classifies the implemented rebalancing strategies, while Section 3.3 presents the test design. Section 3.4 reports the results of our history-based simulation analysis and Section 3.5 discusses potential driving forces that could explain our simulation results. The paper concludes in Section 3.6 by pointing out possible implications for portfolio management.

3.2 Implemented Rebalancing Strategies

Two different types of rebalancing have to be distinguished in the investment practice: periodic and interval rebalancing. While periodic rebalancing demands a reallocation to the predetermined target weights at the end of each period, interval rebalancing requires the implementation of a no-trade region around those target weights. In this study, we concentrate on a combination of both strategies: periodic rebalancing with the additional option of a symmetric no-trade region around the target weights.

Table I presents a classification of all rebalancing strategies under investigation: periodic, threshold, and range rebalancing with yearly, quarterly, and monthly trading intervals. In case of interval rebalancing, the portfolio manager must further differentiate between a reallocation to the target weights (threshold rebalancing) and to the nearest edge of the target weights (range rebalancing). While it is evident that the implementation of a no-trade region reduces transaction costs as a consequence of a reduced portfolio turnover, our analysis sheds light on whether the additional utility of the reduced transaction costs will exceed the utility of a modified risk-return profile. In order to make our results comparable with the investment practice, we mimic the long-term strategy of the GPFG by implementing a symmetric no-trade region of ±3% around the target weights (Norwegian Ministry of Finance (2012)).[1]

[1] Basically, the long-term strategy of the GPFG can be characterized by a quarterly trading frequency, an

Table I – Classification of Implemented Rebalancing Strategies

This table presents all rebalancing strategies under investigation. The periodic rebalancing strategies 1, 2, and 3 are characterized by a regular reallocation to the predetermined target weights at the end of each period. Strategies 4, 5, and 6 represent threshold rebalancing, which is classified as periodic interval rebalancing with a strict adjustment to the target weights. In contrast, the range rebalancing strategies 7, 8, and 9 require a reallocation to the nearest edge of the predefined interval boundaries. A threshold of ±3% is applied to both threshold rebalancing and range rebalancing.

Rebalancing Strategy	Frequency	Threshold	Reallocation	Classification	No.
Yearly Periodic Rebalancing	Yearly	No Threshold	Target Weights	Periodic	1
Quarterly Periodic Rebalancing	Quarterly	No Threshold	Target Weights	Periodic	2
Monthly Periodic Rebalancing	Monthly	No Threshold	Target Weights	Periodic	3
Yearly Threshold Rebalancing	Yearly	Threshold	Target Weights	Threshold	4
Quarterly Threshold Rebalancing	Quarterly	Threshold	Target Weights	Threshold	5
Monthly Threshold Rebalancing	Monthly	Threshold	Target Weights	Threshold	6
Yearly Range Rebalancing	Yearly	Threshold	Interval Boundaries	Range	7
Quarterly Range Rebalancing	Quarterly	Threshold	Interval Boundaries	Range	8
Monthly Range Rebalancing	Monthly	Threshold	Interval Boundaries	Range	9

We explain the exact procedure of the two different classes of interval rebalancing with the help of an example. Assume an investor of a two-asset class portfolio with a strategic asset allocation of 60% stocks and 40% government bonds. Further assume a quarterly rebalancing frequency and a no-trade region of ±3% around the target weights. Conducting a '3% quarterly threshold rebalancing strategy' requires a rebalancing back to the target weights of 60% stocks whenever the relative market capitalization of stocks has moved outside the no-trade region of (57%; 63%) at the end of each quarter. Otherwise, no transactions take place. Supervising a '3% quarterly range rebalancing strategy', the portfolio manager again has to check at the end of each quarter whether the relative market capitalization of stocks has fallen under 57% or has risen above 63%. However, in the first case, a rebalancing to the lower threshold of 57% is necessary, whereas in the second case the relative stock market capitalization has to be adjusted to the upper threshold of 63% by the portfolio manager. If the stocks relative market capitalization lies between (57%; 63%) at the end of the quarter, again no transactions will take place.

implemented no-trade region of ±3% around the target weights, and a reallocation back to the target weights if the relative stock proportion of stocks has fallen outside the no-trade region for one day during the corresponding quarter (Norwegian Ministry of Finance (2012)). This strategy is comparable to the implemented '3% quarterly threshold rebalancing' (strategy 5 in Table I).

3.3 Methodology

3.3.1 Data

In order to conduct a statistical test, a sufficient number of observations is necessary. Ranging from January 1982 to December 2011, our sample period constitutes a reasonable trade-off between the availability of the time series and the number of countries to be included. Based on monthly return data of well-diversified stock and government bond market total return indices as well as money market rates from Thomson Datastream, our analysis comprises the financial markets of the United States, the United Kingdom, and Germany. Moreover, to appropriately reflect real-world practice, the maturities of the government bond time series, namely 5, 7, and 10 years, also determine the corresponding investment horizons to be analyzed. All featuring a maturity of 3 months, we apply Treasury bills (United States), LIBOR (United Kingdom), and FIBOR (Germany) as proxies for the risk-free rate.

Table II presents the descriptive statistics of our dataset. As shown in Panel A of Table II, there are substantial differences between the capital markets of the United States, the United Kingdom, and Germany during the underlying 30-year sample period. For example, the German stock market exhibits the lowest average return with a value of 8.75%, while it also features the highest volatility with a value 22.06%. In contrast, the average stock market return for the United Kingdom is 10.84% with a volatility of 16.14%. Dividing the 30-year sample period of the United States financial markets' into two disjunctive 15-year sub-periods, Panel B of Table II further illustrates that the time series characteristics themselves can change over time. In particular, this is obvious for the average returns of all three asset classes. The stock market return has decreased from 15.59% in the first 15-year subsample to 5.31% in the second one, while the government bond market return has decreased from 10.66% to 6.48%, and the cash market return from 6.21% to 2.71%. As all these time series characteristics will have an impact on the performance of rebalancing, not only an analysis of each country is necessary, but also a methodological approach that allows to preserve most of the time series characteristics and financial market dependencies.

3.3.2 Settings

Evaluating the performance of different rebalancing strategies, we focus on a 60% stocks and 40% government bonds asset allocation for three different reasons. First of all, our analysis considers three different countries (United States, United Kingdom, and Germany), three different classes of rebalancing (periodic, threshold, and range rebalancing), three different investment horizons (5, 7, and 10 year), and three dif-

Table II – Descriptive Statistics

Panel A presents the cross-sectional descriptive statistics of the stock, government bond, and money markets of the United States, the United Kingdom, and Germany over the entire 30-year sample period from January 1982 to December 2011. Panel B shows the descriptive statistics of the United States over the entire 30-year sample period as well as the two corresponding disjunctive 15-year subsamples. Bonds denote government bonds with a maturity of 10 years. Cash represents 3-month money market rates. All statistics are calculated on a monthly basis using continuous compounded returns. Mean, Volatility, Skewness, and Kurtosis denote the annualized mean return, volatility, skewness, and kurtosis. Minimum and Maximum are the monthly minimum and maximum returns.

Panel A: Cross-Sectional Descriptive Statistics

Asset	Statistics	United States	United Kingdom	Germany
Stocks	Mean (%)	10.45	10.84	8.75
	Volatility (%)	15.77	16.14	22.06
	Skewness	-0.91	-1.15	-0.92
	Kurtosis	6.07	8.05	5.60
	Minimum (%)	-23.85	-30.02	-28.67
	Maximum (%)	12.47	13.72	19.02
Bonds	Mean (%)	8.57	10.19	7.34
	Volatility (%)	7.91	8.01	5.53
	Skewness	0.05	-0.06	-0.29
	Kurtosis	3.66	4.45	3.26
	Minimum (%)	-7.36	-8.16	-5.69
	Maximum (%)	9.40	8.17	5.37
Cash (level)	Mean (%)	4.46	6.91	4.43
	Volatility (%)	0.77	1.01	0.65
	Skewness	0.16	0.23	0.55
	Kurtosis	2.70	2.39	2.62
	Minimum (%)	0.00	0.00	0.00
	Maximum (%)	0.01	0.01	0.01

Panel B: Descriptive Statistics of the United States for Subsamples

Asset	Statistics	Full Sample Jan-82 - Dec-11	1st Half Jan-82 - Dec-96	2nd Half Jan-97 - Dec-11
Stocks	Mean (%)	10.45	15.59	5.31
	Volatility (%)	15.77	14.47	16.89
	Skewness	-0.91	-1.12	-0.71
	Kurtosis	6.07	9.79	3.94
	Minimum (%)	-23.85	-23.85	-18.76
	Maximum (%)	12.47	12.47	10.42
Bonds	Mean (%)	8.57	10.66	6.48
	Volatility (%)	7.91	8.10	7.68
	Skewness	0.05	0.08	-0.01
	Kurtosis	3.66	2.85	4.63
	Minimum (%)	-7.36	-4.50	-7.36
	Maximum (%)	9.40	7.30	9.40
Cash (level)	Mean (%)	4.46	6.21	2.71
	Volatility (%)	0.77	0.60	0.57
	Skewness	0.16	0.45	-0.01
	Kurtosis	2.70	3.07	1.44
	Minimum (%)	0.00	0.00	0.00
	Maximum (%)	0.01	0.01	0.01

ferent risk-adjusted performance measures (Sharpe ratio, Sortino ratio, and Omega measure). As all these parameters are linked by multiplication, we thus fix the initial asset allocation at 60% stocks and 40% government bonds in order to concentrate on our primary research question. Secondly, representing one of the world's largest institutional investors by the end of 2011, the GPFG is a predominant example of having pursued a 60% stocks and 40% government bonds asset allocation in the past, thereby reflecting the high relevance for investment practice (Chambers et al. (2012), Norwegian Ministry of Finance (2012)). Thirdly, a 60/40 asset allocation also enables us to compare and discuss our empirical results with the findings of prior rebalancing studies.

Moreover, we incorporate realistic transaction costs quoted at 15 bps per round-trip. In particular, applying well-diversified stock market as well as government bond total return indices, we quote 10 bps for buying/selling stocks and 5 bps for buying/selling bonds.

3.3.3 Motivation

The primary objective of our analysis is the statistical comparison of the performance of different rebalancing strategies under realistic market conditions. For this reason, our simulation approach is based on historical data. The implementation of the stationary bootstrap of Politis and Romano (1994) further enables us to preserve time series characteristics and financial market dependencies (such as positive autocorrelation in the short-run, heteroscedasticity, fat tails, left-skewed return distributions, and asset class correlations) to the greatest possible extent and to derive valuable recommendations for portfolio management.

So far, Monte Carlo simulations have been a suitable approach to analyze the impact of different market conditions on the performance of rebalancing strategies (Jones and Stine (2010), Sun et al. (2006), Donohue and Yip (2003), and Buetow et al. (2002)). However, the rising frequency of financial crises provides strong evidence that commonly used probability distribution functions, such as the normal distribution or the t-distribution, seem to be no longer appropriate for modeling financial markets. Not only the descriptive statistics of Table II substantiates this observation, but also Annaert et al. (2009), among many others, who document that financial return series tend to be left-skewed and exhibit fat tails as well as heteroscedasticity. Moreover, as illustrated by Panel B of Table II, time series characteristics must not be necessarily stable over time. For example, the studies of Ferson et al. (1987) as well as Ferson and Harvey (1991, 1993) all contribute to the explanation of time-varying risk premia. Engle (1982), Engle et al. (1987), and Dumas and Solnik (1995) provide analyses on time-varying risk, and Erb et al. (1994), Ball and Torous (2000), Longin and Solnik (2001), and Buraschi et al.

(2010) examine time-varying asset class correlations. In summary, all these findings indicate that an appropriate calibration of the parameters for a Monte Carlo simulation can be extremely difficult.

Analyses based on real world data avoid most of these difficulties. However, rebalancing studies that conduct simple historical analyses suffer from the drawback of examining only a single realization or a fairly small number of realizations. For example, Jaconetti et al. (2010) analyze an 84-year sample period from 1926 to 2009, Harjoto and Jones (2006) a 10-year sample period from 1995 to 2004, and Tsai (2001) a 15-year sample period from 1986 to 2000. Overall, Jaconetti et al. (2010) as well as Tsai (2001) both argue that there is no universally optimal rebalancing strategy. While Tsai (2001) concludes that it does not matter much which strategy is adopted, Jaconetti et al. (2010) recommend a semiannual or annual 5% threshold rebalancing strategy. Moreover, Harjoto and Jones (2006) report that a 15% monthly threshold rebalancing strategy is superior compared to other rebalancing strategies during all market phases. This issue of mixed results is reinforced by the fact that the performance of rebalancing is highly path-dependent because it constitutes a dynamic portfolio strategy. Therefore, it cannot be ruled out that the empirical results are also driven to a large extent by distinctive features of the underlying sample period rather than by the rebalancing strategy under investigation. Brock et al. (1992) report that this danger of data snooping can be severe, and thus the empirical findings of simple historical analyses do not allow reliable interpretations.

3.3.4 Test Design

Based on the limitations of both Monte Carlo simulations and historical analyses, we perform history-based simulations. Representing a reasonable trade-off, this approach enables us not only to capture most of the time series information, but also to conduct a statistical test, thereby clearly separating our analysis of the performance of rebalancing strategies from both Monte Carlo simulations and historical analyses. In particular, we apply the stationary bootstrap of Politis and Romano (1994) and test whether the mean of a difference time series is equal to zero:

$$H_0 : \Delta_{PM} = 0 \qquad \text{versus} \qquad H_1 : \Delta_{PM} \neq 0, \tag{1}$$

where PM constitutes the risk-adjusted performance measure, which is either the Sharpe ratio, the Sortino ratio, or the Omega measure. The difference between the two performance measures is given by:

$$\Delta_{PM} = PM_A - PM_B, \tag{2}$$

where A and B denote rebalancing strategies as classified in Table I. The arithmetic mean is an appropriate point estimator of (2):

$$\widehat{\Delta}_{PM} = \widehat{PM}_A - \widehat{PM}_B. \tag{3}$$

Before executing our tests, the parameters *performance measure*, *rebalancing class*, *trading frequency*, *country*, and *investment horizon* have to be specified. In a first step, we compare different rebalancing strategies within a given rebalancing class. As classified in Table I, this could either be periodic rebalancing, threshold rebalancing, or range rebalancing. After having determined the risk-adjusted performance measure of interest, we end up with three comparisons for each rebalancing class, each country, and each investment horizon:

Monthly rebalancing - buy-and-hold (M-BAH), (4.1)
Quarterly rebalancing - buy-and-hold (Q-BAH), (4.2)
Yearly rebalancing - buy-and-hold (Y-BAH). (4.3)

Having identified the optimal rebalancing strategy within each rebalancing class, we compare the performance differences between these three rebalancing strategies in a second step:

Periodic rebalancing - threshold rebalancing, (4.4)
Periodic rebalancing - range rebalancing, (4.5)
Threshold rebalancing - range rebalancing. (4.6)

By construction, all these strategies will be very similar in their performance, as already reported in Arnott and Lovell (1993) and Tsai (2001). This similarity comes at the price that it will be difficult to detect any differences in performance measures. In fact, the more similar the corresponding rebalancing strategies under investigation, the more difficult it becomes to uncover statistically significant differences. Therefore, we expect that statistical significance is less pronounced for comparisons of rebalancing strategies which either belong to the same rebalancing class (periodic, threshold, or range rebalancing) or exhibit an identical trading frequency (yearly, quarterly, or monthly). In contrast, we further hypothesize that statistical significance will be more pronounced when we compare the optimal rebalancing strategy of each rebalancing class with each other pairwise. Finally, we conjecture that statistical significance should be most pronounced for comparisons between excessive rebalancing (periodic monthly rebalancing) and very infrequent rebalancing (yearly range rebalancing).

We implement the stationary bootstrap of Politis and Romano (1994). Being applicable to stationary, weakly dependent data, the stationary bootstrap allows us to efficiently exploit time series information by simulating realistic market conditions. In order to generate return paths that could have been realized in the past by drawing blocks of different lengths, we have to ascertain the probability P for resampling the return observations. Following a geometric distribution, we take advantage of the resulting inverse relationship between P and the average block size, which is the expected reciprocal value of P. We determine the optimal average block size by using the automatic block-length selection for the dependent bootstrap of Politis and White (2004), and we further incorporate the corrections made by Patton et al. (2009). Taken as a whole, an average block length of 2 is recommended for all stock and government bond time series of all three countries under investigation. Although Table II reports substantial cross-country differences, this finding allows us to compare our empirical results derived from the different financial markets of the United States, the United Kingdom, and Germany.

We start our analysis by bootstrapping pairwise 100 return paths of stocks, government bonds, and risk-free rates for each country under investigation. The pairwise resampling is necessary in order to preserve the cross-sectional dependency structure between stocks, government bonds, and risk-free-rates. The investment horizons to be analyzed, namely 5, 7, and 10 years, determine the length of the resampled return paths. In contrast to Ledoit and Wolf (2008, 2011) who examine a 10-year investment horizon by bootstrapping from a 10-year sample period, we resample investment horizons of 5, 7, and 10 years by drawing blocks of different lengths from the underlying 30-year sample period. This procedure enables us both to exploit the full information of the underlying sample period and to compare the impact of different investment horizons on the performance of rebalancing. In order to conduct statistical comparisons according to (4.1) – (4.6), we ascertain the rebalancing class, the trading frequency, and the performance measure of interest and calculate the mean for the corresponding difference time series. In a second step, we repeat this procedure B times in order to construct two-sided percentile intervals according to Efron and Tibshirani (1998):

$$\widehat{\Delta}^*_{PM[1]} \leq \widehat{\Delta}^*_{PM[2]} \leq \ldots \leq \widehat{\Delta}^*_{PM[B-1]} \leq \widehat{\Delta}^*_{PM[B]}, \tag{5}$$

where (5) states the ordered difference series of the performance measure of interest. In this context, Romano and Wolf (2006) document that the studentized block bootstrap leads to an improved coverage accuracy for small sample sizes in comparison to normal theory intervals as well as the basic bootstrap. In case of small to moderate sample sizes, Ledoit and Wolf (2008, 2011) also suggest a studentized time series bootstrap if p-values need to be calculated. Nevertheless, covering 30 years with 360 monthly return observations, our sample period can be considered as large, legitimating the construction of percentile intervals as described by Efron and Tibshirani (1998):

$$CI = \left[\widehat{\Delta}^{*}_{PM\left[\frac{\alpha}{2}\cdot B\right]}, \widehat{\Delta}^{*}_{PM\left[\left(1-\frac{\alpha}{2}\right)\cdot B\right]} \right]. \tag{6}$$

The null hypothesis H_0 is rejected at the significance level α if $0 \notin CI$. The nominal levels of α to be considered are 0.01, 0.05, and 0.10. We conduct $B = 1,000$ simulations. Repeated simulations reveal that our results are stable in capturing the underlying sample patterns.

3.4 Empirical Simulation Results

Taking both the return and the risk of a portfolio strategy into account, we apply the Sharpe ratio, the Sortino ratio, and the Omega measure in order to appropriately evaluate portfolio performance. We start our discussion by comparing the risk-adjusted performance on a statistical basis within each of the three rebalancing classes periodic, threshold, and range rebalancing. Our analysis proceeds with the statistical comparison of the risk-adjusted performance of rebalancing between these classes.

3.4.1 Periodic Rebalancing

The Sharpe ratio (Sharpe (1966)) is the most commonly used risk-adjusted performance measure in investment practice. Panel A of Table III shows the average annualized Sharpe ratios of periodic rebalancing classified by trading frequency, investment horizon, and country. On average, quarterly periodic rebalancing exhibits higher Sharpe ratios for all countries and all investment horizons under investigation compared to both monthly and yearly rebalancing. This finding provides a first hint that both too frequent as well as too infrequent rebalancing could lead to an inferior risk-adjusted portfolio performance.

Although we observe a similar pattern for all three countries under investigation, Table III also clearly illustrates the cross-country differences. Assuming a 10-year investment horizon and yearly periodic rebalancing by way of example, the average Sharpe ratio of the United States is 0.579, whereas the average Sharpe ratios of the United Kingdom and Germany are substantially lower with values of 0.389 and 0.355, respectively. Again, classified by country and investment horizon, Table IV proves whether the differences in the average risk-adjusted performance reported in Table III are statistically significant or whether they can simply be ascribed to a distinctive feature of the underlying sample period. If both boundaries are positive (negative), the first rebalancing strategy causes a significantly higher (lower) average risk-adjusted performance compared to the second one. Otherwise, the confidence interval includes

Table III – Average Risk-Adjusted Performance of Periodic Rebalancing
Classified by country and investment horizon, this table shows the average risk-adjusted performance of periodic rebalancing with yearly, quarterly, and monthly trading intervals over the sample period from January 1982 to December 2011. Panel A reports the average annualized Sharpe ratios, Panel B the average annualized Sortino ratios, and Panel C the average Omega measures. All strategies are based on a 60% stocks and 40% bonds asset allocation with a threshold of 0%. Transaction costs are quoted at 15 bps per round-trip. 1,000 simulations with an average block length of 2 are performed. Repeated simulations reveal that the results are stable.

Period	Rebalancing Strategy	United States	United Kingdom	Germany
	Panel A: Average Annualized Sharpe Ratios			
5	Yearly Rebalancing	0.580	0.354	0.354
5	Quarterly Rebalancing	0.583	0.356	0.359
5	Monthly Rebalancing	0.580	0.355	0.356
7	Yearly Rebalancing	0.597	0.398	0.369
7	Quarterly Rebalancing	0.598	0.399	0.372
7	Monthly Rebalancing	0.594	0.398	0.368
10	Yearly Rebalancing	0.579	0.389	0.355
10	Quarterly Rebalancing	0.579	0.390	0.356
10	Monthly Rebalancing	0.575	0.389	0.351
	Panel B: Average Annualized Sortino Ratios			
5	Yearly Rebalancing	2.008	1.927	1.227
5	Quarterly Rebalancing	2.027	1.942	1.246
5	Monthly Rebalancing	2.022	1.941	1.242
7	Yearly Rebalancing	1.945	1.939	1.191
7	Quarterly Rebalancing	1.956	1.949	1.201
7	Monthly Rebalancing	1.949	1.946	1.193
10	Yearly Rebalancing	1.831	1.816	1.116
10	Quarterly Rebalancing	1.836	1.821	1.122
10	Monthly Rebalancing	1.828	1.817	1.112
	Panel C: Average Omega Measures			
5	Yearly Rebalancing	1.380	1.328	0.843
5	Quarterly Rebalancing	1.393	1.340	0.856
5	Monthly Rebalancing	1.390	1.339	0.853
7	Yearly Rebalancing	1.334	1.331	0.815
7	Quarterly Rebalancing	1.340	1.338	0.822
7	Monthly Rebalancing	1.334	1.336	0.817
10	Yearly Rebalancing	1.240	1.234	0.760
10	Quarterly Rebalancing	1.243	1.237	0.764
10	Monthly Rebalancing	1.237	1.234	0.758

zero, implying that the difference is lost in estimation error and that no statistical inferences can be drawn. In eight out of nine cases, monthly periodic rebalancing leads to a significantly lower Sharpe ratio compared to quarterly periodic rebalancing. Although we cannot uncover statistical significance for the financial market of the United Kingdom with an underlying 5-year investment horizon, the position of the 10% quantile (indicated by the magnitude of the lower and upper boundary) suggests that quarterly periodic rebalancing seems to produce a superior risk-adjusted performance in terms of average Sharpe ratios in comparison to monthly periodic rebalancing as well. Even if – in most cases – no significance can be detected, the magnitude of the lower and the upper boundary of the underlying 10% quantiles in Panel A of Table IV further reveals that yearly periodic rebalancing also tends to produce inferior risk-adjusted average Sharpe ratios in comparison to quarterly periodic rebalancing.

From an economic perspective, positive deviations from the target return are not expected to be perceived as risk by investors, but rather as an opportunity to generate an extra return on the invested capital. Therefore, in addition to the Sharpe ratio, we also apply the Sortino ratio, which only takes negative deviations from the expected return into account (Sortino and Price (1994)):

$$S_i(\tau) = \frac{\bar{r}_i - \tau}{\sqrt{\int_{-\infty}^{\tau} (\tau - r_i)^2 f(r_i) dr_i}}, \tag{7}$$

where $\bar{r}_i$ is the average return of the underlying rebalancing strategy i, $f(r_i)$ the corresponding probability density function, and τ the target return required by the investor. We set the target return to zero, which allows us to differentiate between realized gains and losses.

Table IV – CIs: Average Risk-Adjusted Performance of Periodic Rebalancing

Classified by country and investment horizon, this table shows the confidence intervals of the difference time series of periodic rebalancing of the average Sharpe ratio (Panel A), of the average Omega measure (Panel B), and of the average Sortino ratio (Panel C). The sample period ranges from January 1982 to December 2011. All strategies are based on a 60% stocks and 40% bonds asset allocation with a threshold of 0%. Transaction costs are quoted at 15 bps per round-trip. Y denotes yearly periodic rebalancing, Q quarterly periodic rebalancing, and M monthly periodic rebalancing. For example, M-Q denotes the difference time series of 'Monthly periodic rebalancing minus quarterly periodic rebalancing'. For each two strategies that are compared, the lower and upper boundary of the confidence interval is calculated. 1,000 simulations with an average block length of 2 are performed. Repeated simulations reveal that the results are stable. *, **, and *** represent statistical significance at the 10%, 5%, and 1% level, respectively. If there is no statistical significance, the corresponding 10% quantiles are reported.

Period	Strategies	United States		United Kingdom		Germany	
Panel A: Average Annualized Sharpe Ratio of Periodic Rebalancing							
5	M-Q	−0.0059	−0.0004***	−0.0025	0.0002	−0.0058	0.0000**
5	M-Y	−0.0042	0.0038	−0.0016	0.0044	−0.0024	0.0074
5	Q-Y	−0.0005	0.0065	0.0000	0.0053*	0.0010	0.0101*
7	M-Q	−0.0065	−0.0015***	−0.0028	−0.0001*	−0.0078	−0.0010***
7	M-Y	−0.0062	0.0011	−0.0027	0.0030	−0.0060	0.0032
7	Q-Y	−0.0019	0.0044	−0.0009	0.0042	−0.0013	0.0071
10	M-Q	−0.0066	−0.0020***	−0.0029	−0.0003**	−0.0081	−0.0022***
10	M-Y	−0.0072	−0.0005*	−0.0035	0.0015	−0.0077	0.0003
10	Q-Y	−0.0028	0.0032	−0.0016	0.0028	−0.0026	0.0049

continued

Table IV – *Continued*

Panel B: Average Annualized Sortino Ratio of Periodic Rebalancing							
5	M-Q	−0.0100	0.0008	−0.0060	0.0028	−0.0102	0.0014
5	M-Y	−0.0000	0.0278	0.0018	0.0271**	0.0015	0.0285*
5	Q-Y	0.0003	0.0389***	0.0011	0.0316***	0.0042	0.0338**
7	M-Q	−0.0126	−0.0014**	−0.0073	0.0009	−0.0141	0.0020**
7	M-Y	−0.0086	0.0144	−0.0031	0.0168	−0.0086	0.0147
7	Q-Y	−0.0000	0.0206	0.0010	0.0186*	0.0004	0.0221*
10	M-Q	−0.0149	−0.0015***	−0.0068	−0.0003*	−0.0169	0.0034***
10	M-Y	−0.0120	0.0076	−0.0067	0.0094	−0.0137	0.0055
10	Q-Y	−0.0031	0.0152	−0.0022	0.0124	−0.0032	0.0148
Panel C: Average Omega Measure of Periodic Rebalancing							
5	M-Q	−0.0068	0.0194	−0.0037	0.0164	−0.0057	0.0205
5	M-Y	0.0006	0.0194*	0.0022	0.0201**	0.0006	0.0215**
5	Q-Y	0.0012	0.0274***	0.0013	0.0224***	0.0004	0.0253***
7	M-Q	−0.0108	−0.0011***	−0.0047	0.0008	−0.0085	−0.0010**
7	M-Y	−0.0070	0.0079	−0.0016	0.0114	−0.0048	0.0100
7	Q-Y	−0.0002	0.0130	0.0012	0.0126*	0.0007	0.0142*
10	M-Q	−0.0110	−0.0027***	−0.0055	−0.0004**	−0.0103	−0.0021***
10	M-Y	−0.0095	0.0029	−0.0052	0.0050	−0.0082	0.0040
10	Q-Y	−0.0025	0.0086	−0.0018	0.0074	−0.0018	0.0097

Although the economic impact seems to be small, Panel B of Table III substantiates the observation that there may be an optimal trading frequency. Again, quarterly periodic rebalancing produces the highest risk-adjusted portfolio performance for all countries and all investment horizons compared to both monthly and yearly periodic rebalancing. All in all, Panel B of Table IV reconfirms our findings in Panel A of Table IV. In five out of nine cases, quarterly periodic rebalancing leads to significantly higher average Sortino ratios compared to both monthly and yearly periodic rebalancing. With regard to the remaining cases, the positions of the 10% quantile also indicate without any exception that quarterly periodic rebalancing tends to exhibit a higher risk-adjusted portfolio performance in terms of average Sortino ratios.

As neither the Sharpe ratio nor the Sortino ratio account for higher moments, such as the skewness of a return distribution or its kurtosis, portfolio recommendations derived on the basis of these risk-adjusted performance measures could be biased. By way of example, since the turn of the millennium, the dot.com bubble burst of 2000, the destabilization effects of 9/11, the subprime mortgage crisis of 2007, and – most recently – the European sovereign debt crisis of 2010 all have impressively shown that fat tails must not be ignored. For this reason, we additionally use the Omega measure, which considers the entire return distribution (Shadwick and Keating (2002)). Representing a special case of the more general performance measure Kappa (Kaplan and Knowles (2004)), it is defined as the ratio of gains to losses relative to a predefined target return:

$$\Omega_i(\tau) = \frac{\int_{\tau}^{\infty}(1 - F(r_i))dr_i}{\int_{-\infty}^{\tau} F(r_i)dr_i}, \tag{8}$$

where $F(r_i)$ denotes the cumulative distribution function of the monthly return r of rebalancing strategy i, and τ is the investor's required rate of return, which we again set to zero. The results in Panel C of Table III and in Panel C of Table IV are qualitatively similar to those in Panel B of Table III and in Panel B of Table IV, thereby substantiating the empirical finding that quarterly periodic rebalancing tends to exhibit a superior risk-adjusted performance compared to both monthly and yearly periodic rebalancing.

Taken as a whole, we conclude that both too frequent as well as too infrequent rebalancing results in a suboptimal portfolio performance. Following Ledoit and Wolf (2011), we also modify the average block length to six instead of two as a further robustness check. The results (not reported) are even stronger in this case. In fact, we even observe that monthly periodic rebalancing is significantly outperformed in comparison to quarterly periodic rebalancing for all three countries, all three investment horizons, and all three risk-adjusted performance measures under investigation. Therefore, our results contradict the recommendation of Arnott and Lovell (1993), who suggest that investors with a long investment horizon should rebalance on a monthly basis.

3.4.2 Interval Rebalancing

Searching for an optimal rebalancing strategy, we additionally test threshold and range rebalancing strategies by implementing a symmetric no-trade region around the target weights. Once a rebalancing threshold is introduced, there are two cases that need to be distinguished with regard to the practical implementation. If a rebalancing is necessary at the end of the predetermined period, the portfolio weights can be reallocated either to the original target weights (Buetow et al. (2002), Harjoto and Jones (2006)) or to the nearest edge of the original target weights (Leland (1999)). As explained above, threshold rebalancing refers to strategies (4) – (6) whereas range rebalancing corresponds to strategies (7) – (9) in Table I.

Table V provides only weak statistical evidence that quarterly threshold rebalancing leads to a better risk-adjusted performance compared to both monthly and yearly threshold rebalancing. However, the positions of the 10% quantile again indicate that quarterly threshold rebalancing seems to produce the highest average risk-adjusted performance for all countries under investigation. This pattern changes with respect to range rebalancing. Table VI documents that yearly range rebalancing is signifi-

Table V – CIs: Average Risk-Adjusted Performance of Threshold Rebalancing

Classified by country and investment horizon, this table shows the confidence intervals of the difference time series of threshold rebalancing of the average Sharpe ratio (Panel A), of the average Omega measure (Panel B), and of the average Sortino ratio (Panel C). The sample period ranges from January 1982 to December 2011. All strategies are based on a 60% stocks and 40% bonds asset allocation with a threshold of 3%. Transaction costs are quoted at 15 bps per round-trip. Y denotes yearly threshold rebalancing, Q quarterly threshold rebalancing, and M monthly threshold rebalancing. For example, M-Q denotes the difference time series of 'Monthly threshold rebalancing minus quarterly threshold rebalancing'. For each two strategies that are compared, the lower and upper boundary of the confidence interval is calculated. 1,000 simulations with an average block length of 2 are performed. Repeated simulations reveal that the results are stable. *, **, and *** represent statistical significance at the 10%, 5%, and 1% level, respectively. If there is no statistical significance, the corresponding 10% quantiles are reported.

Period	Strategies	United States		United Kingdom		Germany	
Panel A: Average Annualized Sharpe Ratio of Periodic Rebalancing							
5	M-Q	−0.0032	0.0014	−0.0021	0.0015	−0.0043	0.0012
5	M-Y	−0.0022	0.0058	−0.0007	0.0051	−0.0013	0.0083
5	Q-Y	−0.0011	0.0065	−0.0002	0.0053	0.0000	0.0107**
7	M-Q	−0.0036	0.0006	−0.0024	0.0012	−0.0049	−0.0002*
7	M-Y	−0.0038	0.0034	−0.0021	0.0038	−0.0045	0.0046
7	Q-Y	−0.0021	0.0045	−0.0011	0.0041	−0.0016	0.0069
10	M-Q	−0.0038	0.0000*	−0.0025	0.0005	−0.0066	−0.0003***
10	M-Y	−0.0049	0.0018	−0.0028	0.0024	−0.0061	0.0018
10	Q-Y	−0.0027	0.0035	−0.0018	0.0031	−0.0026	0.0050
Panel B: Average Annualized Sortino Ratio of Threshold Rebalancing							
5	M-Q	−0.0075	0.0083	−0.0055	0.0071	−0.0092	0.0053
5	M-Y	0.0004	0.0286*	0.0006	0.0254**	0.0011	0.0279*
5	Q-Y	0.0010	0.0277*	0.0002	0.0248**	0.0016	0.0313**
7	M-Q	−0.0083	0.0055	−0.0070	0.0052	−0.0106	0.0017
7	M-Y	−0.0054	0.0180	−0.0049	0.0161	−0.0076	0.0167
7	Q-Y	−0.0030	0.0192	−0.0022	0.0163	−0.0021	0.0202
10	M-Q	−0.0093	0.0029	−0.0079	−0.0008**	−0.0118	0.0029
10	M-Y	−0.0089	0.0112	−0.0073	0.0096	−0.0120	0.0075
10	Q-Y	−0.0056	0.0138	−0.0053	0.0115	−0.0049	0.0135
Panel C: Average Omega Measure of Threshold Rebalancing							
5	M-Q	−0.0053	0.0053	−0.0036	0.0051	−0.0053	0.0033
5	M-Y	0.0002	0.0187*	0.0003	0.0180**	0.0007	0.0187*
5	Q-Y	0.0009	0.0183*	0.0005	0.0172**	0.0017	0.0201**
7	M-Q	−0.0063	0.0028	−0.0047	0.0033	−0.0065	0.0009
7	M-Y	−0.0048	0.0102	−0.0033	0.0105	−0.0049	0.0106
7	Q-Y	−0.0027	0.0117	−0.0016	0.0107	−0.0013	0.0126
10	M-Q	−0.0064	0.0010	−0.0047	0.0014	−0.0078	−0.0006**
10	M-Y	−0.0069	0.0058	−0.0055	0.0053	−0.0076	0.0048
10	Q-Y	−0.0039	0.0079	−0.0036	0.0067	−0.0031	0.0083

cantly outperformed by both quarterly and monthly range rebalancing in almost all cases. Moreover, monthly range rebalancing also tends to feature a better risk-adjusted performance compared to quarterly range rebalancing.

Table VI – CIs: Average Risk-Adjusted Performance of Range Rebalancing

Classified by investment horizon, this table shows the confidence intervals of the difference time series of range rebalancing of the average Sharpe ratio (Panel A), of the average Omega measure (Panel B), and of the average Sortino ratio (Panel C). The sample period ranges from January 1982 to December 2011. All strategies are based on a 60% stocks and 40% bonds asset allocation with a threshold of 3%. Transaction costs are quoted at 15 bps per round-trip. Y denotes yearly range rebalancing, Q quarterly range rebalancing, and M monthly range rebalancing. For example, M-Q denotes the difference time series of 'Monthly range rebalancing minus quarterly range rebalancing'. For each two strategies that are compared, the lower and upper boundary of the confidence interval is calculated. 1,000 simulations with an average block length of 2 are performed. Repeated simulations reveal that the results are stable. *, **, and *** represent statistical significance at the 10%, 5%, and 1% level, respectively. If there is no statistical significance, the corresponding 10% quantiles are reported.

Period	Strategies	United States		United Kingdom		Germany	
		Panel A: Average Annualized Sharpe Ratio of Range Rebalancing					
5	M-Q	−0.0002	0.0021	0.0000	0.0018*	0.0002	0.0034*
5	M-Y	0.0018	0.0099***	0.0014	0.0076***	0.0038	0.0159***
5	Q-Y	0.0017	0.0084***	0.0009	0.0065***	0.0026	0.0133***
7	M-Q	−0.0008	0.0013	−0.0004	0.0013	−0.0007	0.0022
7	M-Y	0.0007	0.0071**	0.0005	0.0056**	0.0006	0.0125***
7	Q-Y	0.0000	0.0069***	0.0000	0.0059***	0.0005	0.0110***
10	M-Q	−0.0012	0.0008	−0.0008	0.0007	−0.0012	0.0013
10	M-Y	−0.0001	0.0049	−0.0004	0.0036	0.0001	0.0081**
10	Q-Y	0.0001	0.0051**	0.0000	0.0034*	0.0005	0.0076**
		Panel B: Average Annualized Sortino Ratio of Range Rebalancing					
5	M-Q	0.0013	0.0105**	0.0010	0.0099***	0.0012	0.0111**
5	M-Y	0.0139	0.0444***	0.0092	0.0359***	0.0135	0.0472***
5	Q-Y	0.0102	0.0380***	0.0062	0.0287***	0.0091	0.0375***
7	M-Q	−0.0001	0.0068	0.0000	0.0061*	−0.0006	0.0070
7	M-Y	0.0058	0.0358***	0.0024	0.0297***	0.0041	0.0369***
7	Q-Y	0.0054	0.0296***	0.0021	0.0243***	0.0037	0.0303***
10	M-Q	−0.0014	0.0045	−0.0015	0.0034	−0.0022	0.0041
10	M-Y	0.0016	0.0264***	0.0007	0.0170**	0.0002	0.0264***
10	Q-Y	0.0025	0.0232***	0.0010	0.0148**	0.0014	0.0251***
		Panel C: Average Omega Measure of Range Rebalancing					
5	M-Q	0.0007	0.0070**	0.0006	0.0071***	0.0010	0.0073**
5	M-Y	0.0090	0.0323***	0.0065	0.0257***	0.0083	0.0317***
5	Q-Y	0.0071	0.0268***	0.0046	0.0200***	0.0060	0.0252***
7	M-Q	−0.0007	0.0040	−0.0001	0.0042	−0.0003	0.0045
7	M-Y	0.0033	0.0218***	0.0010	0.0205***	0.0024	0.0235***
7	Q-Y	0.0033	0.0194***	0.0009	0.0164***	0.0025	0.0205***
10	M-Q	−0.0011	0.0027	−0.0013	0.0020	−0.0013	0.0026
10	M-Y	0.0008	0.0169***	0.0004	0.0096*	0.0001	0.0172***
10	Q-Y	0.0012	0.0146***	0.0003	0.0086*	0.0009	0.0157***

Overall, our results provide strong evidence that both excessive rebalancing (monthly periodic rebalancing) as well as too infrequent rebalancing (yearly range rebalancing) leads on average to inferior Sharpe ratios, Sortino ratios, and Omega measures. As a consequence, our empirical findings indicate that there may be an optimal rebalancing strategy. However, this result contradicts the reasoning of Jaconetti et al. (2010) as well as that of Tsai (2001), who all conclude from their analyses that there is no universally optimal rebalancing strategy. In contrast, we show that the optimal trading patterns change with respect to the underlying rebalancing strategy (periodic, threshold or range rebalancing). While a quarterly trading frequency seems optimal for periodic and threshold rebalancing, it is a monthly trading frequency that tends to produce the best results for range rebalancing.

3.4.3 Optimal Rebalancing

Having analyzed the trading patterns within a given rebalancing class, the question of interest that now arises is which of these rebalancing strategies performs best. Therefore, we compare the average risk-adjusted performance of quarterly periodic rebalancing, quarterly threshold rebalancing, and monthly range rebalancing with each other. Table VII documents that monthly range rebalancing leads to a significantly lower risk-adjusted performance in terms of Sharpe ratios, Sortino ratios, and Omega measures at least at the 5% level compared with both quarterly periodic rebalancing and quarterly threshold rebalancing. Although we cannot detect statistical significance with regard to average Sharpe ratios, Panel B as well as Panel C of Table VII document that, on average, quarterly periodic rebalancing offers significantly higher Sortino ratios and Omega measures for all three countries and all investment horizons under investigation compared to quarterly threshold rebalancing.

As expected, statistical significance is more pronounced for pairwise comparisons between the optimal rebalancing strategies of each rebalancing class in contrast to pairwise comparisons within a particular rebalancing class. Although we do not present all possible pairwise comparisons of rebalancing strategies for the sake of brevity, we observe statistical significance at least at the 5% level in terms of the Omega measure for all comparisons between quarterly periodic rebalancing and quarterly threshold rebalancing. As statistical significance is less pronounced for pairwise comparisons within a specific rebalancing class, as shown by Tables IV, V, and VI, respectively, we note that the rebalancing algorithm itself will have a higher impact on the performance of rebalancing compared to the trading frequency.

In conclusion, if an investor identifies rebalancing as an appropriate portfolio strategy subject to the underlying risk and return preferences as well as the regulatory environment, our empirical simulation analysis supports a quarterly periodic rebalancing strategy. This finding is in line with the long-term strategy of the GPFG, which also adopts a quarterly trading frequency (Chambers et al. (2012), Norwegian Ministry of Finance (2012)). However, we further show that the benefit from reduced transaction

Table VII – CIs: Average Risk-Adjusted Performance of Optimal Rebalancing

Classified by investment horizon, this table shows the confidence intervals of the difference time series of the average annualized Sharpe ratio between quarterly periodic rebalancing, quarterly threshold rebalancing, and monthly range rebalancing. The sample period ranges from January 1982 to December 2011. All strategies are based on a 60% stocks and 40% bonds asset allocation. The no-trade region comprises ±3% around the target weights. Transaction costs are quoted at 15 bps per round-trip. For example, Periodic-Range denotes the difference time series of 'Quarterly periodic rebalancing minus monthly range rebalancing'. For each two strategies that are compared, the lower and upper boundary of the confidence interval is calculated. 1,000 simulations with an average block length of 2 are performed. Repeated simulations reveal that the results are stable. *, **, and *** represent statistical significance at the 10%, 5%, and 1% level, respectively. If there is no statistical significance, the corresponding 10% quantiles are reported.

Period	Strategies	United States		United Kingdom		Germany	
				Panel A: Average Annualized Sharpe Ratio			
5	Periodic-Threshold	−0.0004	0.0036	−0.0006	0.0025	−0.0007	0.0029
5	Periodic-Range	0.0024	0.0102***	0.0010	0.0069***	0.0010	0.0088***
5	Threshold-Range	0.0007	0.0079***	0.0001	0.0058***	0.0003	0.0079***
7	Periodic-Threshold	−0.0007	0.0029	−0.0005	0.0024	−0.0008	0.0024
7	Periodic-Range	0.0010	0.0082***	0.0008	0.0068***	0.0005	0.0076***
7	Threshold-Range	0.0002	0.0065***	0.0006	0.0047**	0.0003	0.0059**
10	Periodic-Threshold	−0.0008	0.0021	−0.0007	0.0018	−0.0008	0.0019
10	Periodic-Range	0.0000	0.0063***	0.0004	0.0045**	0.0006	0.0053**
10	Threshold-Range	0.0003	0.0048**	0.0001	0.0037**	0.0002	0.0047**
				Panel B: Average Annualized Sortino Ratio			
5	Periodic-Threshold	0.0026	0.0197**	0.0014	0.0191***	0.0008	0.0109*
5	Periodic-Range	0.0179	0.0472***	0.0166	0.0414***	0.0088	0.0303***
5	Threshold-Range	0.0078	0.0349***	0.0071	0.0292***	0.0035	0.0240***
7	Periodic-Threshold	0.0010	0.0154**	0.0000	0.0174***	0.0003	0.0090*
7	Periodic-Range	0.0019	0.0360***	0.0113	0.0338***	0.0058	0.0253***
7	Threshold-Range	0.0049	0.0268***	0.0051	0.0240***	0.0018	0.0211***
10	Periodic-Threshold	0.0004	0.0114**	0.0006	0.0113**	0.0002	0.0070*
10	Periodic-Range	0.0077	0.0263***	0.0062	0.0252***	0.0042	0.0192***
10	Threshold-Range	0.0029	0.0211***	0.0019	0.0177***	0.0008	0.0162***
				Panel C: Average Omega Measure			
5	Periodic-Threshold	0.0010	0.0159***	0.0020	0.0146***	0.0005	0.0080**
5	Periodic-Range	0.0127	0.0346***	0.0127	0.0315***	0.0067	0.0208***
5	Threshold-Range	0.0055	0.0249***	0.0057	0.0213***	0.0033	0.0160***
7	Periodic-Threshold	0.0012	0.0101**	0.0012	0.0122***	0.0005	0.0065**
7	Periodic-Range	0.0088	0.0244***	0.0096	0.0250***	0.0051	0.0174***
7	Threshold-Range	0.0036	0.0181***	0.0046	0.0172***	0.0021	0.0142***
10	Periodic-Threshold	0.0004	0.0072**	0.0003	0.0085***	0.0003	0.0051**
10	Periodic-Range	0.0051	0.0171***	0.0056	0.0177***	0.0037	0.0136***
10	Threshold-Range	0.0022	0.0138***	0.0021	0.0122***	0.0014	0.0109***

costs due to the implemented no-trade region around the target weights does not outweigh a lower value added resulting from altered risk and return characteristics. With respect to the GPFG, two additional arguments are worth noting that advantage quarterly periodic rebalancing in comparison to quarterly threshold rebalancing. First of all, transaction costs of the GPFG are expected to be lower than 15 bps per round-trip due to its bargaining power. Secondly, our analysis is based on lump-sum payments taking place at the beginning of the underlying investment horizon, whereas the GPFG receives a steady cash inflow from selling a part of Norway's petroleum resources. Reducing the need for reallocating the portfolio weights, this partial rebalancing also contributes to saving transaction costs.

Although the economic impact seems to be small at a first glance, it is an important finding for investment practice that our primary results are statistically significant and robust across countries. Even small differences are expected to be economically relevant if AuM are of considerable size. With over 550 billion US$ AuM by the end of 2011, the GPFG is a good example for a large institutional investor who conducts rebalancing as a cost-efficient rule-based investment strategy. In order to illustrate and compare the return potential between monthly and quarterly periodic rebalancing, we construct a hypothetical example similar to the GPFG. Obtaining the input parameters from the financial market of the United States, we assume an average Sharpe ratio of 0.575 for monthly periodic rebalancing and 0.579 for quarterly periodic rebalancing, respectively. We further keep the underlying volatility constant at 10% p.a.[2] The resulting implicit annual excess return would be 5.75% for monthly periodic rebalancing and 5.79% for quarterly periodic rebalancing. Therefore, given that both strategies exhibit the same risk potential, the overall excess return potential would amount to 220 million US$.

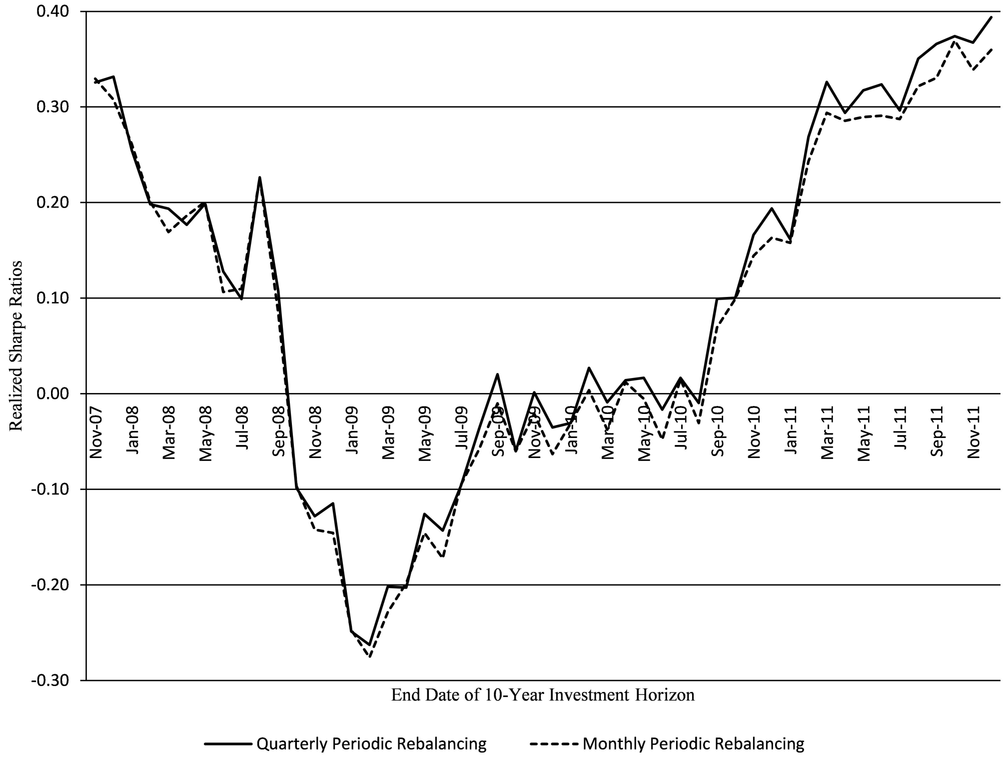
Taking the closest investor perspective possible, we further substantiate our history-based simulation findings by illustrating the risk-adjusted performance of both quarterly and monthly periodic rebalancing of past investment periods. In particular, we apply a rolling window approach, which enables us to calculate the corresponding Sharpe ratios of overlapping 10-year investment horizons realized in the past. In order to completely capture the impact of the 2007 subprime mortgage crisis, we start our analysis in January 1998, as December 2007 officially represents the beginning of the resulting economic downturn according to the National Bureau of Economic Research (2012). Therefore, the last observation of the first 10-year investment horizon of our rolling window approach is December 2007.

Illustrating the resulting Sharpe ratios of both quarterly and monthly periodic rebalancing by way of example, Panel A of Figure I confirms our history-based simulation results. Two observations are particularly noteworthy. Firstly, the historical Sharpe ratios of both strategies are very close to each other, thus confirming our simulation

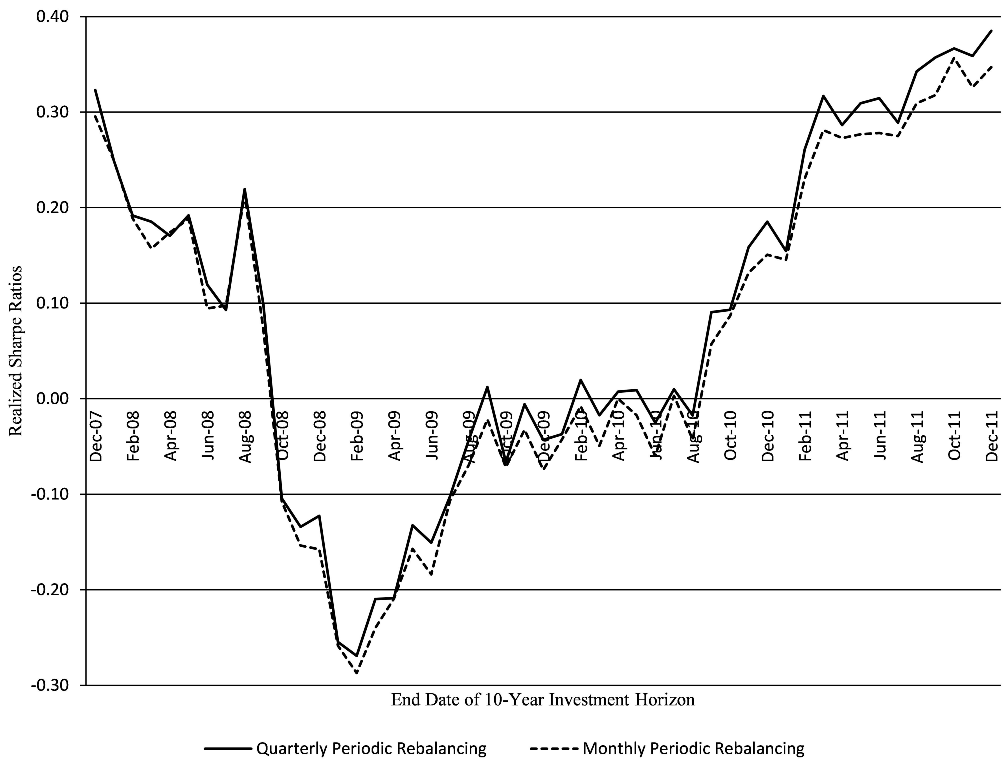
---

[2] In results not shown, monthly periodic rebalancing exhibits – on average – a marginally lower annual return as well as a marginally higher annual risk in terms of volatility and semi-volatility for all countries and all investment horizons compared to quarterly periodic rebalancing, leading to inferior average Sharpe ratios.

Figure I – Sharpe Ratios of a 10-Year Rolling Window



(A) 15 bps Transaction Costs



(B) 100 bps Transaction Costs

findings where the differences in performance measures are marginal. This similarity of different rebalancing strategies is also documented in Arnott and Lovell (1993) as well as Tsai (2001). Secondly, quarterly periodic rebalancing again clearly outperforms monthly periodic rebalancing. This effect even increases with higher transaction costs. In order to stay focused on our main contribution – the statistical comparison of different rebalancing strategies' risk-adjusted performance – this study focuses on developed stock and governments bond markets with a high trading volume. Therefore, quoting transaction costs at 15 bps per round-trip can be seen as rather conservative. However, many institutional investors also invest in less liquid markets, which causes higher transaction costs. Therefore, Panel B of Figure I quotes transaction costs at 100 bps per round-trip (Pesaran and Timmermann (1994)) in order to show the impact of higher transaction costs on the risk-adjusted performance of rebalancing. Transaction costs have a negative impact on the level of risk-adjusted performance of both quarterly and monthly periodic rebalancing. However, it is apparent that this effect is only of minor importance. Moreover, the higher the trading frequency, the more pronounced the negative impact of higher transaction costs will be. As expected, Panel B of Figure I shows that the outperformance of quarterly periodic rebalancing compared to monthly periodic rebalancing is stronger with higher transaction costs – even though this effect is marginal.

Overall, the rolling window approach represents actual investor performance and confirms the results of our history-based simulation approach. As quarterly and monthly periodic rebalancing differ only in the trading intervals, the difference in risk-adjusted performance is very close in most instances. However, although transaction costs are quoted at *only* 15 bps per round-trip, both our history-based simulation and our rolling window approach provide strong evidence that, on average, quarterly periodic rebalancing outperforms monthly periodic rebalancing.

3.4.4 Impact of Time Series Characteristics on Portfolio Performance

A still open question is which driving force is responsible for the observation that, on average, monthly periodic rebalancing exhibits a lower risk-adjusted performance compared to quarterly periodic rebalancing. This result remains valid even if transaction costs are excluded from our analysis (not reported). Possible sources could be time series characteristics, such as short-term momentum and long-term mean-reversion, the cross-correlation between stocks and bonds, and distributional characteristics of the return generating process. In order to shed light on this issue, we conduct a simple Monte Carlo simulation, assuming a geometric Brownian motion with normally distributed stock and government bond markets returns as well as a correlation of zero in a first step. We further calibrate the parameters 'mean' and 'volatility' by applying the average values of the US financial markets over the entire 30-year sample period. Based on Table II, we use 10.45% (8.57%) as the average annual sample mean for stocks (government bonds) and 15.77% (7.91%) as the corresponding average annual volatility. Furthermore, we keep the risk-free rate constant at the long-term average of 4.46%. Our simulations (results not tabulated) do not detect any differences in risk-adjusted performance measures between monthly and quarterly periodic rebalancing.

In a second step, we completely break down the time series structure of our real-world data by using a fixed block length of 1 in the stationary bootstrap. This procedure deletes short-term momentum, but preserves both the distributional characteristics as well as the correlation structure between stocks and bonds. Again, we do not detect any statistical differences in risk-adjusted performance measures between monthly and quarterly periodic rebalancing (results not tabulated). As statistical significance completely disappears if the time series structure is destroyed, we conclude that time series characteristics – especially short-term momentum – are the primary sources capable of explaining the statistically significant differences in average Sharpe ratios between monthly and quarterly periodic rebalancing. While we do not have a simple explanation at hand, we suspect that the interrelations between returns, risk, and – in particular – portfolio weights are responsible for this finding.

3.5 Conclusion

This study compares the risk-adjusted performance of different rebalancing strategies under realistic market conditions by reporting statistical significance levels. First of all, we document that monthly periodic rebalancing features a lower average risk-adjusted performance for all three countries and for all investment horizons under investigation in comparison to quarterly periodic rebalancing. Moreover, as yearly range rebalancing also leads to inferior Sharpe ratios, Sortino ratios, and Omega measures, our results imply that there is an optimal rebalancing strategy with both excessive rebalancing (monthly periodic rebalancing) as well as too infrequent rebalancing (yearly range rebalancing) provoking a suboptimal risk-adjusted performance. Secondly, the optimal trading frequency is subject to the underlying rebalancing algorithm. Within the corresponding rebalancing class, quarterly periodic, quarterly threshold, and monthly range rebalancing seem to produce the highest risk-adjusted performance. Thirdly, as quarterly periodic rebalancing leads to significantly higher average Sortino ratios and Omega ratios in comparison to quarterly threshold and monthly range rebalancing, our history-based simulation promotes a quarterly periodic rebalancing strategy as it delivers the highest average risk-adjusted performance for all three countries under investigation. As a robustness test, our rolling window approach underpins this finding. Investigating realized 10-year investment horizons, quarterly periodic rebalancing outperforms monthly periodic rebalancing by way of example. Lastly, short-term momentum seems to be the primary source capable of explaining the statistically significant differences between monthly and quarterly periodic rebalancing.

References

Annaert, J., S. Van Osselaer, and B. Verstraete (2009). “Performance Evaluation of Portfolio Insurance Strategies Using Stochastic Dominance Criteria.” In: *Journal of Banking and Finance* 33, pp. 272–280.

Arnott, R.D. and R.M. Lovell (1993). “Rebalancing: Why? When? How Often?” In: *Journal of Investing* 2, pp. 5–10.

Ball, C. and W.N. Torous (2000). “Stochastic Correlation Across International Stock Markets.” In: *Journal of Empirical Finance* 7, pp. 373–388.

Brock, W., J. Lakonishok, and B. LeBaron (1992). “Simple Technical Trading Rules and the Stochastic Properties of Stock Returns.” In: *Journal of Finance* 47, pp. 1731–1764.

Buetow, G.W., E. Hunt, R. Sellers, D. Trotter, and W.A. Whipple (2002). “The Benefits of Rebalancing.” In: *Journal of Portfolio Management* 28, pp. 23–32.

Buraschi, A., P. Porchia, and F. Trojani (2010). “Correlation Risk and Optimal Portfolio Choice.” In: *Journal of Finance* 65, pp. 393–420.

Chambers, D., E. Dimson, and A. Ilmanen (2012). “The Norway Model.” In: *Journal of Portfolio Management* 38, pp. 67–81.

Donohue, C. and K. Yip (2003). “Optimal Portfolio Rebalancing with Transaction Costs.” In: *Journal of Portfolio Management* 29, pp. 49–63.

Dumas, B. and B. Solnik (1995). “The World Price of Foreign Exchange Risk.” In: *Journal of Finance* 50, pp. 445–479.

Efron, B. and R.J. Tibshirani (1998). *An Introduction to the Bootstrap*. Chapman & Hall/CRC.

Engle, R.F. (1982). “Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation.” In: *Econometrica* 50, pp. 987–1007.

Engle, R.F., D.M. Lilien, and R.P. Robins (1987). "Estimating Time-Varying Risk Premia in the Term Structure: The ARCH-M Model." In: *Econometrica* 55, pp. 391–408.

Erb, C.B., C.R. Harvey, and T.E. Viskanta (1994). "Forecasting International Equity Correlations." In: *Financial Analysts Journal* 50, pp. 32–45.

Fabozzi, F.J., S.M. Focardi, and P.N. Kolm (2006). "Incorporating Trading Strategies in the Black-Litterman Framework." In: *Journal of Trading* 1, pp. 28–37.

Ferson, W.E. and C.R. Harvey (1991). "The Variation of Economic Risk Premiums." In: *Journal of Political Economy* 99, pp. 385–415.

– (1993). "The Risk and Predictability of International Equity Returns." In: *Review of Financial Studies* 6, pp. 527–566.

Ferson, W.E., S. Kandel, and R.F. Stambaugh (1987). "Tests of Asset Pricing with Time-Varying Expected Risk Premiums and Market Betas." In: *Journal of Finance* 42, pp. 201–220.

Harjoto, M.A. and F.J. Jones (2006). "Rebalancing Strategies for Stocks and Bonds Asset Allocation." In: *Journal of Wealth Management* 9, pp. 37–44.

Jaconetti, C.M., F.M. Kinniry, and Y. Zilbering (2010). *Best Practices for Portfolio Rebalancing*. Tech. rep. Vanguard.

Jones, S.K. and J.B. Stine (2010). "Expected Utility and the Non-Normal Returns of Common Portfolio Rebalancing Strategies." In: *Journal of Asset Management* 10, pp. 406–419.

Kaplan, P.D. and J.A. Knowles (2004). "Kappa: A Generalized Downside Risk-Adjusted Performance Measure." In: *Journal of Performance Measurement* 8, pp. 42–54.

Ledoit, O. and M. Wolf (2008). "Robust Performance Hypothesis Testing with the Sharpe Ratio." In: *Journal of Empirical Finance* 15, pp. 850–859.

– (2011). "Robust Performance Hypothesis Testing with Variance." In: *Wilmott Magazine* 55, pp. 86–89.

Leibowitz, M.L. and A. Bova (2011). "Policy Portfolios and Rebalancing Behavior." In: *Journal of Portfolio Management* 37, pp. 60–71.

Leland, H. E. (1999). "Optimal Portfolio Management with Transactions Costs and Capital Gains Taxes." Research Program in Finance Working Paper RPF-290, University of California, Berkeley, 1999.

Longin, F. and B. Solnik (2001). “Extreme Correlation of International Equity Markets.” In: *Journal of Finance* 56, pp. 649–676.

National Bureau of Economic Research (2012). *US Business Cycle Expansions and Contractions*. URL: http://www.nber.org/cycles/cyclesmain.html..

Norwegian Ministry of Finance (2012). *Report No. 17 to the Storting: The Management of the Government Pension Fund in 2011*. URL: http://www.regjeringen.no/pages/37635499/stm_gpf.pdf.

Patton, A., D.N. Politis, and H. White (2009). “Correction to “Automatic Block-Length Selection for the Dependent Bootstrap” by D. Politis and H. White.” In: *Econometric Reviews* 28, pp. 372–375.

Perold, A.F. and W.F. Sharpe (1988). “Dynamic Strategies for Asset Allocation.” In: *Financial Analysts Journal* 44, pp. 16–27.

Pesaran, M.H. and A. Timmermann (1994). “Forecasting Stock Returns: An Examination of Stock Market Trading in the Presence of Transaction Costs.” In: *Journal of Forecasting* 13, pp. 335–367.

Politis, D.N. and J.P. Romano (1994). “The Stationary Bootstrap.” In: *Journal of the American Statistical Association* 89, pp. 1303–1313.

Politis, D.N. and H. White (2004). “Automatic Block-Length Selection for the Dependent Bootstrap.” In: *Econometric Reviews* 23, pp. 53–70.

Romano, J.P. and M. Wolf (2006). “Improved Nonparametric Confidence Intervals in Time Series Regressions.” In: *Journal of Nonparametric Statistics* 18, pp. 199–214.

Shadwick, W.F. and C. Keating (2002). “A Universal Performance Measure.” In: *Journal of Performance Measurement* 6, pp. 59–84.

Sharpe, W.F. (1966). “Mutual Fund Performance.” In: *Journal of Business* 39, pp. 119–138.

Sortino, F.A. and L.N. Price (1994). “Performance Measurement in a Downside Risk Framework.” In: *Journal of Investing* 3, pp. 59–64.

Sun, W., A.C. Fan, L.-W. Chen, T. Schouwenaars, and M.A. Albota (2006). “Optimal Rebalancing for Institutional Portfolios.” In: *Journal of Portfolio Management* 32, pp. 33–43.

Swensen, D. F. (2009). *Pioneering Portfolio Management: An Unconventional Approach to Institutional Investment*. New York: The Free Press.

The Yale Endowment (2012). *Endowment Update 2012*. URL: http://investments.yale.edu/images/documents/yale_endowment_12.pdf.

Tsai, C.S.-Y. (2001). "Rebalancing Diversified Portfolios of Various Risk Profiles." In: *Journal of Financial Planning* 14, pp. 104–110.

Chapter 4

Testing Rebalancing Strategies for Stock-Bond Portfolios Across Different Asset Allocations

Abstract

We compare the risk-adjusted performance of stock-bond portfolios between rebalancing and buy-and-hold across different asset allocations by reporting statistical significance levels. Our investigation is based on a 30-year dataset and incorporates the financial markets of the United States, the United Kingdom, and Germany. In order to draw reasonable recommendations to portfolio management, we conduct a history-based simulation approach which enables us to simulate realistic market conditions. Our empirical results show that a frequent rebalancing significantly enhances risk-adjusted portfolio performance for all stock-bond portfolios if the stocks' portfolio weight exceeds a certain threshold, depending on the country and on the risk-adjusted performance measure under investigation.

4.1 Introduction

Strategic asset allocation constitutes an integral component of portfolio performance. Analyzing the impact of strategic asset allocation on portfolio performance without removing the effects of market movements, Brinson et al. (1986), Brinson et al. (1991), as well as Ibbotson and Kaplan (2000) all provide evidence that 100% of the return level and about 90% of the variation of a fund's returns over time is explained by its investment policy. While the first result states an identity in aggregate market conditions derived from the fact that active management must be a zero sum game by definition, the second result does not differentiate between the effect of overall market movements and the impact of strategic asset allocation on the variation of a fund's return. Accounting for the impact of market movements, Xiong et al. (2010) document that strategic asset allocation and active management are equally important for explaining return differences on aggregate. Although strategic asset allocation is less important than predicted by prior studies, the findings of Xiong et al. (2010) reconfirm the rationale that the investment policy is of considerable importance to investment practice.

Therefore, building on the fundamentals of modern portfolio theory by Markowitz (1952), many optimization routines have been developed by both academic researchers and practitioners. The primary objective is the construction of reasonable portfolios that meet investors' different requirements. Developments based on the mean-variance portfolio optimization of Markowitz (1952) are still the best known and most widely used portfolio optimization approaches in practice. However, the implementation of mean-variance optimizers also raises several problems. Michaud (1989) as well as Kritzman (2006) report that mean-variance optimization routines tend to under- and overweight assets or even entire asset classes with the highest input errors, often resulting in a non-optimal risk-adjusted performance. Moreover, if assets are close substitutes for one another, small errors of the input parameters could result in significant misallocations of the portfolio. For these reasons, other more robust optimization routines have been developed for practice.[1] Instead of focusing on these portfolio optimization techniques, this paper analyzes the risk-adjusted performance of stock-bond portfolios by systematically changing the underlying portfolio weights. Moreover, given this initial asset allocation, we further investigate whether a frequent rebalancing contributes to significantly enhancing risk-adjusted portfolio performance.

Although this question is of utmost importance for investment practice, only Tsai (2001) examines the impact of rebalancing on portfolio performance with a focus on different risk profiles. In particular, she investigates the Sharpe ratio of five different stock-bond portfolios over the sample period from January 1986 to December 2000. Considering seven asset classes, the equity portfolio weight varies between 20%, 40%, 60%, 80%, and 98%, respectively. While it is not obvious which specific rebalancing algorithm should be adopted, all four rebalancing strategies under investigation produce a higher Sharpe ratio for all five portfolios in comparison to buy-and-hold. However, it must not be ignored that this analysis is based on one single 15-year period, which intensifies the potential problem of data snooping. It could be the case that the empirical results are substantially driven by specific characteristics of the underlying sample period while the investment strategy under investigation could have only a minor impact on the risk-adjusted portfolio performance (Brock et al. (1992)).

In order to shed light on this issue, we compare the risk-adjusted performance of a given asset allocation between buy-and-hold and rebalancing strategies on a statistical basis. In particular, we apply the stationary bootstrap of Politis and Romano (1994). Preserving most of the time series characteristics, this history-based simulation approach enables us to report statistical significance levels under realistic market conditions.

Our results provide evidence that rebalancing significantly outperforms buy-and-hold if the portfolio weight of stocks exceeds a certain threshold. Depending on the country under investigation and on the risk-adjusted performance measure, this threshold ranges between 0% and 30%. We further document that the optimal asset allocation is subject to both the country and the period under investigation.

The remainder of this article is structured as follows: While Section 4.2 briefly describes the data and provides descriptive statistics, Section 4.3 classifies the implemented rebalancing strategies. We proceed by presenting the analyzed asset allocations in Section 4.4 and explaining the applied simulation set-up in Section 4.5. Section 4.6 discusses the empirical results and Section 4.7 concludes the article.

[1] See: Black and Litterman (1992), Fabozzi et al. (2007), DeMiguel, Garlappi, and Uppal (2009b), and DeMiguel, Garlappi, Nogales, and Uppal (2009a).

4.2 Data and Descriptive Statistics

In order to test whether the empirical findings are robust across countries, our analysis not only comprises the financial market of the United States, but also those of the United Kingdom and of Germany. The monthly data covers the sample period from January 1982 to December 2011. We obtain well-diversified stock and government bond total return indices as well as Treasury bills (United States), LIBOR (United Kingdom), and FIBOR (Germany) from Thomson Reuters Datastream. As the Treasury bills, the LIBOR, and the FIBOR represent liquid instruments featuring a high trading volume on the secondary market, they serve as proxies for the risk-free rates which are necessary to calculate the Sharpe ratio. They all exhibit a maturity of three months.

Table I – Descriptive Statistics

This table presents the cross-sectional descriptive statistics of the stock, government bond, and money markets of the United States, the United Kingdom, and Germany over the entire 30-year sample period from January 1982 to December 2011. Bonds denote government bonds with a maturity of 10 years. Cash represents the corresponding 3-month money market rates. All statistics are calculated on a monthly basis using continuous compounded returns. Mean, Volatility, Skewness, and Kurtosis denote the annualized mean return, volatility, skewness, and kurtosis. Skewness and Kurtosis are calculated as the third and fourth normalized centered moments. Minimum and Maximum are the monthly minimum and maximum returns.

Asset	Statistics	United States	United Kingdom	Germany
Stocks	Mean (%)	10.45	10.84	8.75
	Volatility (%)	15.77	16.14	22.06
	Skewness	-0.91	-1.15	-0.92
	Kurtosis	6.07	8.05	5.60
	Minimum (%)	-23.85	-30.02	-28.67
	Maximum (%)	12.47	13.72	19.02
Bonds	Mean (%)	8.57	10.19	7.34
	Volatility (%)	7.91	8.01	5.53
	Skewness	0.05	-0.06	-0.29
	Kurtosis	3.66	4.45	3.26
	Minimum (%)	-7.36	-8.16	-5.69
	Maximum (%)	9.40	8.17	5.37
Cash (level)	Mean (%)	4.46	6.91	4.43
	Volatility (%)	0.77	1.01	0.65
	Skewness	0.16	0.23	0.55
	Kurtosis	2.70	2.39	2.62
	Minimum (%)	0.00	0.00	0.00
	Maximum (%)	0.01	0.01	0.01

The descriptive statistics in Table I show substantial cross-country differences, thereby legitimating the expansion of our analysis to the financial markets of the United Kingdom and of Germany. On the one hand, for example, the German stock market exhibits the lowest average annualized return while simultaneously featuring the highest annualized volatility of all three countries. As a result, both the stock market of the United States as well as that of the United Kingdom offer a better risk-return ratio in comparison to Germany during the underlying 30-year sample period. On the other hand, the German government bond market exhibits the lowest volatility of all three countries under investigation. Thus, we expect that the optimal asset allocation of a German stock-bond portfolio exhibits a higher proportion of government bonds in comparison to optimal stock-bond portfolios of the United States and of the United Kingdom. Furthermore, the low correlation between stocks and government bonds provides a first hint that a pronounced diversification effect will be found. Overall, these cross-country differences stress the need to shed light on the issue whether rebalancing leads to a significantly better risk-adjusted performance compared to buy-and-hold across all three countries under investigation or not.

4.3 Implemented Rebalancing Strategies

As shown in Table II, we focus on ten different rebalancing strategies which are categorized by the underlying rebalancing algorithm into four distinctive rebalancing classes: (i) buy-and-hold, (ii) periodic rebalancing, (iii) threshold rebalancing, and (iv) range rebalancing. If no rebalancing period is defined, rebalancing reduces to buy-and-hold. That is, no transactions take place during the entire investment period until divestment. In contrast, periodic rebalancing is characterized by a reallocation to the initial portfolio weights at the end of each predetermined period. We consider yearly, quarterly, and monthly trading frequencies. In this context, transaction costs are quoted at 15 bps per round-trip throughout the entire analysis. They can be composed in 10 bps for buying/selling stocks and 5 bps for buying/selling government bonds.

In order to reduce portfolio turnover and hence, save transaction costs, a no-trade region around the target weights can be implemented as well. In line with Norway's Government Pension Fund Global, this study focuses on a 3% symmetric no-trade interval around the target weights (Norwegian Ministry of Finance (2012)). Given the strategic decision to conduct interval rebalancing, two different approaches must be distinguished. While threshold rebalancing requires a readjustment to the target weights at the end of each predetermined period, range rebalancing further reduces portfolio turnover by rebalancing the assets back only to the nearest edge of the no-trade interval at the end of each pre-assigned period (Leland (1999)). However, the increased utility of saved transaction costs must be opposed to the modified risk-return characteristics in order to draw valuable recommendations to portfolio management. Our history-based simulation takes both aspects into account.

The exact rebalancing procedure is explained with the help of an example: Consider a portfolio consisting of 60% stocks and 40% government bonds as well as a yearly

Table II – Classification of Implemented Rebalancing Strategies
This table presents all rebalancing strategies under investigation. The periodic rebalancing strategies 2, 3, and 4 are characterized by a regular reallocation to the predetermined target weights at the end of each period. Strategies 5, 6, and 7 represent threshold rebalancing, which is classified as periodic interval rebalancing with a strict adjustment to the target weights. In contrast, the range rebalancing strategies 8, 9, and 10 require a reallocation to the nearest edge of the predefined interval boundaries. A threshold of ±3% is applied to both threshold rebalancing and range rebalancing.

Rebalancing Strategy	Frequency	Threshold	Reallocation	Classification	No.
Buy-and-Hold	No Adjustments	No Threshold	No Reallocation	Buy-and-Hold	1
Yearly Periodic Rebalancing	Yearly	No Threshold	Target Weights	Periodic	2
Quarterly Periodic Rebalancing	Quarterly	No Threshold	Target Weights	Periodic	3
Monthly Periodic Rebalancing	Monthly	No Threshold	Target Weights	Periodic	4
Yearly Threshold Rebalancing	Yearly	Threshold	Target Weights	Threshold	5
Quarterly Threshold Rebalancing	Quarterly	Threshold	Target Weights	Threshold	6
Monthly Threshold Rebalancing	Monthly	Threshold	Target Weights	Threshold	7
Yearly Range Rebalancing	Yearly	Threshold	Interval Boundaries	Range	8
Quarterly Range Rebalancing	Quarterly	Threshold	Interval Boundaries	Range	9
Monthly Range Rebalancing	Monthly	Threshold	Interval Boundaries	Range	10

trading frequency. A ’3% yearly threshold rebalancing strategy’ requires rebalancing the assets back to the initial 60/40 asset allocation whenever the stocks’ portfolio weight exceeds the no-trade interval $[57\%, 63\%]$ at the end of each year. In contrast, a ’3% yearly range rebalancing strategy’ postulates a readjustment at the end of the year to 57% if the relative proportion of stocks has fallen below 57%, or to 63% if the stocks’ portfolio weight exceeds 63%. Otherwise, the stocks’ portfolio weight lies within the no-trade region, implying that no transactions are required.

4.4 Analyzed Asset Allocations

Prior research provides evidence that rebalancing generates a value added for institutional investors. Examining a 50/50 stock-bond portfolio over the period from 1968 to 1991, Arnott and Lovell (1993) show that almost all analyzed rebalancing strategies exhibit a higher Treynor ratio compared to buy-and-hold. Despite the highest portfolio turnover, they recommend a monthly rebalancing strategy to long-term investors as this strategy features the highest return while its corresponding volatility is only slightly higher compared to those of buy-and-hold. In contrast, the analysis of Harjoto and Jones (2006) is based on a 60/40 stock-bond portfolio and covers the period from 1995 to 2004. During the underlying sample period, buy-and-hold exhibits the lowest Sharpe ratio in comparison to all remaining rebalancing strategies. Although Harjoto and Jones (2006) show the need to rebalance the portfolio, they recommend not to readjust the portfolio weights too frequently. Focusing on a 60/40 stock-bond asset allocation again, Tokat and Wicas (2007) investigate the period from 1960 to 2003. Comparing several rebalancing strategies with buy-and-hold in different market environments, they document that the frequent reallocation to the target weight helps to control risk in all market environments. Jaconetti et al. (2010) reconfirm this finding. Based on a 60% stocks and 40% bonds asset allocation, they investigate the period from 1926 to 2009. During this 84-years sample period, the stocks' portfolio weight of a buy-and-hold strategy is as high as 99% at maximum and as low as 36% at minimum. Hence, Jaconetti et al. (2010) conclude that the primary objective of rebalancing is the reduction of risk relative to a given asset allocation.

However, prior research on rebalancing lacks a focus on investigating different asset allocations. While Arnott and Lovell (1993) consider a 50/50 asset allocation, Harjoto and Jones (2006), Tokat and Wicas (2007) as well as Jaconetti et al. (2010) all concentrate on a 60/40 stock-bond portfolio. Only Tsai (2001) examines the impact of rebalancing on the risk-adjusted performance across different asset allocations. Covering the sample period from 1986 to 2000, she investigates five different risk profiles with an equity proportion of 20%, 40%, 60%, 80%, and 98%, respectively. While it is not clear which specific rebalancing strategy should be adopted, all four strategies under investigation outperform buy-and-hold in terms of Sharpe ratios for all five different risk profiles. Nevertheless, the explanatory power of her analysis is weakened by the fact that it is based on one single 15-year period.

With regard to the portfolio mix, the contribution of our rebalancing study is twofold to academic research and practitioners. First of all, we conduct a systematic analysis of the risk-adjusted performance of all rebalancing strategies shown in Table II across different asset allocations. Focusing on a two-asset class portfolio consisting of stocks and government bonds, we systematically change the relative proportion of stocks in 10% steps from 0% to 100%. Overall, we analyze the risk-adjusted performance of eleven different asset allocations for each country under investigation. Secondly, for each of these eleven different portfolio mixes, we further test whether rebalancing is able to significantly enhance risk-adjusted portfolio performance compared to buy-and-hold.

4.5 Simulation Set-Up

The primary objective of this study is the statistical comparison of the impact between rebalancing and buy-and-hold on the risk-adjusted performance of stock-bond portfolios across different asset allocations.

4.5.1 Resampling Procedure

In order to draw valuable recommendations to investment practice, we conduct a history-based simulation approach. As this procedure is based on historical data, we use the stationary bootstrap of Politis and Romano (1994), which is applicable to stationary, weakly dependent data. In particular, we randomly draw blocks of different lengths with replacement from the entire 30-year sample period.

We illustrate the exact resampling procedure for one single 10-year investment horizon with the help of an example. Imagine two boxes. Representing the monthly return data of the original 30-year sample period, the first box includes 360 numbered balls from 1 to 360. Without any exception, all return observations are randomly drawn with replacement from the first box. The 120 spots-of the simulated 10-year return path are consecutively filled according to the selection process. A distinctive feature of the applied stationary bootstrap of Politis and Romano (1994) constitutes the selection of blocks with different lengths, which allows preserving most of the original time series properties. In order to generate blocks of different length, a second box is necessary. By way of example, it contains one red ball and four green balls.

In a first step, we randomly sample with replacement one single return observation from the first box. Accordingly, the probability of being selected is 1/360 for each single return observation. Let's say, we have randomly drawn ball no. 237. Thus, ball no. 237 is the first return observation of our simulated 10-year return path. In a second step, we now sample one ball with replacement from the second box. If we draw a red ball, we will again start with the first step. However, if we sample a green ball, we will fill the second spot of the 10-year investment horizon with ball no. 238 from the first box. In this case, we proceed by again drawing a ball from the second box. If we receive a red ball, we will again start with the first step, but if we draw a green ball again, we will fill the third spot with ball no. 239 from the first box, and so on. Overall, we randomly sample with replacement from the second box and successively fill the spots of the underlying 10-year investment horizon with consecutive return observations until we have drawn a red ball. In this case, we begin with the first step as described above. We repeat this procedure until all 120 spots have been filled successively. Drawing either a red or green ball from the second box in our example corresponds to a geometrical distribution with an underlying success probability of $p = 20\%$, where the red ball denotes a 'success' and the green ball 'no success', respectively. Finally, we have simulated a single 10-year return path that could have been realized in the past as it is based on historical return data. Not only are the return paths of stocks simulated for all three countries under investigation at identical points in time. Simultaneously, we apply the same procedure to government bonds and risk-free rates. This procedure as a whole preserves the correlation across the three different asset classes and further enables us to compare our results across countries.

Overall, the only parameter left to be determined of the described resampling procedure is the probability P, which constitutes the relevant factor of the different block lengths to be drawn. The higher P is, the shorter the expected block length will be.

However, instead of assigning an appropriate probability to P, we exploit the reciprocal relation between the average block length and the probability P resulting from the geometric distribution. In particular, the expected value of a geometrically distributed random variable is its reciprocal value $1/p$. In order to determine an appropriate average block length, we apply the automatic block-length selection for the dependent bootstrap of Politis and White (2004) as well as the corrections made by Patton et al. (2009). Overall, an average block length of 2 is recommended to both the stock and the government bond market for all three countries under investigation.[2]

4.5.2 Construction of Confidence Intervals

We address the investment practice by constructing confidence intervals that are easy to understand and straightforward to implement. Therefore, we follow the description of Efron and Tibshirani (1998) which lays the foundation of our analysis whether rebalancing leads on average to a higher risk-adjusted performance compared to buy-and-hold. In particular, we test whether the mean of a difference time series is equal to zero:

$$H_0 : \Delta_{PM} = 0 \qquad \text{versus} \qquad H_1 : \Delta_{PM} \neq 0, \tag{1}$$

where PM denotes the risk-adjusted performance measure. In this study, we focus on the Sharpe ratio, the Sortino ratio, and the Omega measure as appropriate risk-adjusted performance measures in order to stay focused on our primary contribution.[3] Furthermore, the rebalancing strategies as well as the country under investigation need to be specified in order to calculate the resulting difference between the risk-adjusted performances of both strategies:

$$\Delta_{PM} = PM_A - PM_B, \tag{2}$$

where A and B constitute rebalancing strategies as shown in Table II. The arithmetic

[2] The implementation of the stationary bootstrap exactly follows the algorithm as decribed by Politis and Romano (1994). This procedure also implies that the data, once the last observation is reached, wraps around to a circle to the starting observation in order to avoid a trapezoidal weithing of the observations at the beginning and at the end of the sample period. For example, consider a resampled block with four consecutive return observations starting at 259. Hence, this block consists of the return observations 259, 360, 1, and 2.

[3] Academic research as well as investment practice analyze a variety of risk-adjusted performance measures. Nevertheless, Eling (2008) finds out that the selection of the performance measure is not critical to performance evaluation. Adcock et al. (2012) further provides evidence that even if returns are not normally distributed, the rank correlation of the Sharpe ratio and other performance measures is one or very close to it.

mean is an appropriate point estimator of (2):

$$\widehat{\Delta}_{PM} = \widehat{PM}_A - \widehat{PM}_B. \tag{3}$$

Again, an example best illustrates the exact implementation of the confidence interval. As described below, we compare the average risk-adjusted performance between quarterly periodic rebalancing and buy-and-hold for the financial market of the United States by employing the Sharpe ratio.

In a first step, we apply both strategies to the simulated 10-year return path and calculate the two corresponding Sharpe ratios. In order to test whether quarterly periodic rebalancing outperforms buy-and-hold on average, we repeat this procedure 100 times in second step and calculate the average Sharpe ratio for both quarterly periodic rebalancing and buy-and-hold. Accordingly, we receive one average Sharpe ratio for each strategy. In a third step, we calculate the difference between these two averages by subtracting the average Sharpe ratio of the buy-and-hold strategy from the average Sharpe ratio of the quarterly periodic rebalancing strategy. However, if this difference is positive, we still do not know whether quarterly periodic rebalancing outperforms buy-and-hold on average. There also is the chance that the true value is negative. Therefore, we derive a distribution of equation (3) by repeating the steps one to three $1,000$ times:

$$\widehat{\Delta}^*_{PM[1]} \leq \widehat{\Delta}^*_{PM[2]} \leq \ldots \leq \widehat{\Delta}^*_{PM[999]} \leq \widehat{\Delta}^*_{PM[1,000]}, \tag{4}$$

where (4) states the ordered difference series of the performance measure of interest. In a last step, we construct percentile intervals as described by Efron and Tibshirani (1998):[4]

$$CI = \left[\widehat{\Delta}^*_{PM\left[\frac{\alpha}{2}\cdot 1,000\right]}, \widehat{\Delta}^*_{PM\left[\left(1-\frac{\alpha}{2}\right)\cdot 1,000\right]}\right]. \tag{5}$$

The nominal level of α to be considered is 0.10. The null hypothesis H_0 is rejected at the 10% level if $0 \notin CI$:

$$CI = \left[\widehat{\Delta}^*_{PM[50]}, \widehat{\Delta}^*_{PM[950]}\right]. \tag{6}$$

[4] Our analysis focuses on a straightforward implementation of confidence intervals. As our sample period covers 360 monthly return observations and thus can be considered a rather large sample, the construction of percentile intervals according to Efron and Tibshirani (1998) is legitimated. Nevertheless, Romano and Wolf (2006) provide evidence that the studentized block bootstrap leads to an improved coverage accuracy for small sample sizes, whereas Ledoit and Wolf (2008, 2011) suggest a studentized time series bootstrap in case of small to moderate sample sizes.

This analysis is based on 100,000 return paths of stocks, government bonds, and risk-free rates. Although these return paths are simulated, we use the same paths for the entire investigation in order to be able to compare our simulation results across different performance measures and countries, as well as a changing asset allocation.

Overall, this history-based simulation approach enables us to preserve most of the time series properties as well as to conduct a systematic comparison by reporting statistical significance levels. Repeated simulations reveal that our results are stable in capturing the underlying sample patterns.

4.6 Empirical Results

Although several studies do report that rebalancing seems to provide a value added to institutional investors, three questions remain unanswered. First of all, there are no studies on rebalancing with a focus on institutional investors outside the United States. However, according to Table I, there are substantial differences in time series characteristics between the United States, the United Kingdom, and Germany. It is not clear whether these cross-country differences have a different impact on the risk-adjusted performance of rebalancing.

Secondly, little research has been done on rebalancing across different asset allocations. While Arnott and Lovell (1993) examine a 50/50 stock-bond portfolio, the analyses of Harjoto and Jones (2006), Tokat and Wicas (2007), and of Jaconetti et al. (2010) all focus on a 60/40 stock-bond portfolio. Only Tsai (2001) investigates the risk-adjusted performance of rebalancing across different asset allocations. However, her analysis only provides a snapshot of the particular state of the market because it is merely based on a single 15-year sample period.

Thirdly, most studies presented above are based on historical analyses. However, as rebalancing constitutes a dynamic trading strategy, its performance is highly path-dependent. Thus, it is possible that the empirical results of those studies are more influenced by specifics of the underlying sample period rather than by the corresponding rebalancing algorithm under investigation. According to Brock et al. (1992), this danger of data snooping could be serious. For example, consider a market environment with a very low volatility and a well-pronounced market trend. Perold and Sharpe (1988) show on a theoretical basis that buy-and-hold must outperform any rebalancing strategy in this particular state of the market because the rebalancing algorithm requires buying past losers and investing the proceeds in past winners. More precisely, the regular reallocation to the worse performing asset not only reduces the upside potential in prolonged upswing markets, but also increases the downside potential in persistent cyclical downturns. This example illustrates that the resulting question of interest must be whether, on average, rebalancing leads to a higher risk-adjusted performance in comparison to buy-and-hold.

By simulating 100,000 different return paths that could have been realized in the past, our history-based simulation approach avoids this potential problem of data snooping. Instead, we are able to conduct a systematic analysis of the risk-adjusted performance of stock-bond portfolios between rebalancing and buy-and-hold by reporting statistical significance levels. In particular, we do not only apply the Sharpe ratio, but also the Sortino ratio and the Omega measure in order to appropriately measure risk-adjusted portfolio performance.

4.6.1 Sharpe Ratio

Focusing on the financial market of the United States by way of example, Panel A of Table III presents the average Sharpe ratios not only of all eleven different asset allocations, but also of all ten rebalancing strategies under investigation. As expected by the low correlation between stocks and government bonds shown in Table I, the diversification effect is well-pronounced.

Table III – Average Sharpe Ratios Across Different Asset Allocations

Panel A presents the average Sharpe ratios of all rebalancing strategies under investigation across different asset allocations for the financial market of the United States. Panel B shows the increase (decrease) of the average Sharpe ratio of the corresponding rebalancing strategy in comparison to the average Sharpe ratio of a buy-and-hold strategy. The sample period ranges from January 1982 to December 2011. Transaction costs are quoted at 15 bps per round-trip. 1,000 simulations with an average block length of 2 are performed. Repeated simulations reveal that the results are stable.

Strategies	Proportion of Stocks										
	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
	Panel A: Average Sharpe Ratios										
BAH	0.523	0.617	0.668	0.671	0.642	0.598	0.552	0.507	0.468	0.435	0.406
Periodic Rebalancing											
Yearly	0.523	0.610	0.669	0.687	0.668	0.627	0.579	0.530	0.484	0.443	0.406
Quarterly	0.523	0.608	0.667	0.686	0.667	0.627	0.579	0.530	0.484	0.443	0.406
Monthly	0.523	0.607	0.664	0.681	0.663	0.623	0.575	0.527	0.482	0.442	0.406
Threshold Rebalancing											
Yearly	0.523	0.610	0.669	0.687	0.668	0.627	0.578	0.529	0.483	0.441	0.406
Quarterly	0.523	0.610	0.669	0.687	0.668	0.627	0.578	0.529	0.483	0.441	0.406
Monthly	0.523	0.609	0.668	0.685	0.666	0.625	0.526	0.528	0.482	0.441	0.406
Range Rebalancing											
Yearly	0.523	0.612	0.670	0.686	0.665	0.623	0.573	0.524	0.479	0.438	0.406
Quarterly	0.523	0.611	0.670	0.687	0.667	0.625	0.576	0.527	0.480	0.438	0.406
Monthly	0.523	0.611	0.670	0.686	0.666	0.625	0.576	0.527	0.480	0.438	0.406

continued

Table III – *Continued*

	Panel B: Average Increase in Sharpe Ratios in %										
Periodic Rebalancing											
Yearly	0.0	−1.2	0.2	2.4	4.0	4.8	4.9	4.4	3.4	1.9	0.0
Quarterly	0.0	−1.4	0.0	2.2	3.9	4.8	5.0	4.5	3.4	1.9	0.0
Monthly	0.0	−1.6	−0.5	1.6	3.2	4.0	4.2	3.8	3.0	1.7	0.0
Threshold Rebalancing											
Yearly	0.0	−1.1	0.3	2.4	4.0	4.7	4.8	4.2	3.1	1.4	0.0
Quarterly	0.0	−1.2	0.2	2.4	4.0	4.8	4.8	4.3	3.2	1.4	0.0
Monthly	0.0	−1.2	0.1	2.1	3.7	4.4	4.5	4.0	3.0	1.4	0.0
Range Rebalancing											
Yearly	0.0	−0.7	0.4	2.2	3.5	4.1	4.0	3.3	2.2	0.7	0.0
Quarterly	0.0	−0.9	0.4	2.4	3.8	4.5	4.4	3.8	2.5	0.8	0.0
Monthly	0.0	−1.0	0.3	2.3	3.8	4.4	4.4	3.8	2.6	0.8	0.0

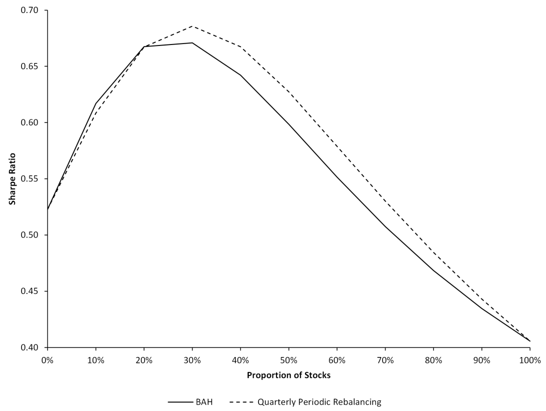
For example, while buy-and-hold leads to an average Sharpe ratio of 0.67 for a 30/70 stock-bond portfolio, the average Sharpe of a 100/0 and a 0/100 stock-bond portfolio accounts only to 0.52 and 0.41, respectively. With regard to the underlying 30-year sample period, an allocation of 30% stocks and 70% government bonds leads to the highest average risk-adjusted portfolio performance for all ten rebalancing strategies. However, in results not reported, the optimal asset allocation strongly depends on the period under investigation.

If the relative proportion of stocks accounts for at least 20%, buy-and-hold is outperformed in terms of Sharpe ratios by all nine rebalancing strategies. Panel B of Table III substantiates this observation by pointing out the average increase in risk-adjusted performance of the respective strategy compared to buy-and-hold. As all three classes (periodic, threshold, and range) and all three trading frequencies (yearly, quarterly, monthly) lead to a higher risk-adjusted performance on average, rebalancing seems to consistently provide a value added to institutional investors across different asset alloctions (if the stock allocation exceeds a certain threshold). However, the question whether these differences are also statistically significant still remains unanswered.

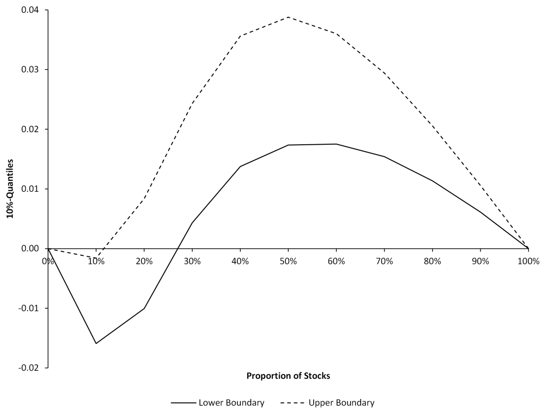
Therefore, Panel A of Figure I shows the average Sharpe ratios of both quarterly periodic rebalancing and buy-and-hold across different asset allocations for the financial market of the United States. Panel B of Figure I presents the corresponding 10%-quantiles. If both boundaries are positive (negative), statistical significance is observed at the 10% level. As all parameters are linked by multiplication, we only illustrate the risk-adjusted performance and the corresponding confidence intervals for buy-and-hold as well as periodic quarterly rebalancing. Given an initial asset allocation of 60% stocks and 40% government bonds, Dichtl et al. (2013) document that quarterly periodic rebalancing provides the highest risk-adjusted performance on average.

With respect to the remaining rebalancing strategies classified in Table II, the results differ only slightly. Comparing the risk-adjusted performance between rebalancing and buy-and-hold, we find a very similar pattern for all rebalancing strategies under investigation.

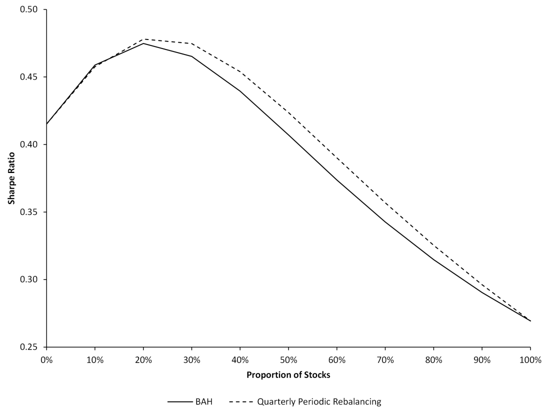
Figure I – Sharpe Ratio Across Different Asset Allocations



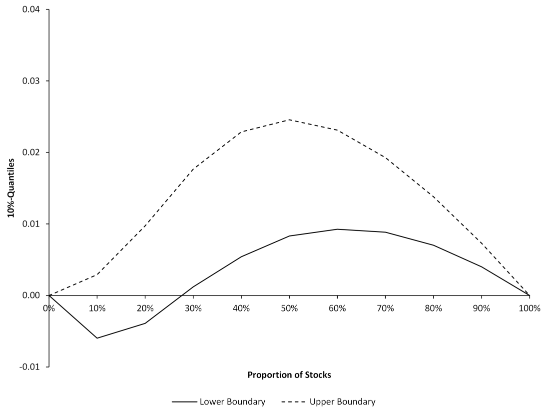
(A) Average Sharpe Ratio (United States)



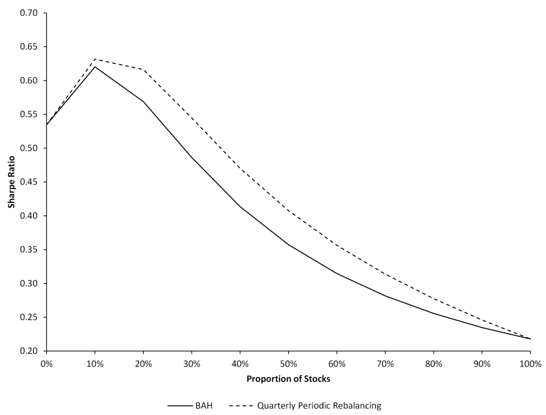
(B) 10%-Quantiles of the Average Sharpe Ratio



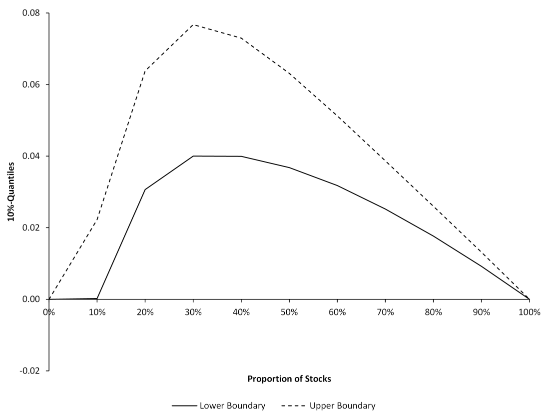
(C) Average Sharpe Ratio (United Kingdom)



(D) 10%-Quantiles of the Average Sharpe Ratio



(E) Average Sharpe Ratio (Germany)



(F) 10%-Quantiles of the Average Sharpe Ratio

As shown by Panel A of Figure I, quarterly periodic rebalancing leads on average to higher Sharpe ratios compared to buy-and-hold if the stocks portfolio weight accounts for at least 30%. Panel B further proves that this finding is statistically significant at the 10% level as both interval boundaries are positive. In contrast, if the relative proportion of stocks falls below 20%, buy-and-hold outperforms quarterly periodic rebalancing. In results not reported, this pattern disappears with both an increasing threshold and an increasing trading frequency – hence the more similar rebalancing becomes to buy-and-hold.

We observe a similar pattern for the financial market of the United Kingdom as shown in Panel C and Panel D of Figure I. Again, quarterly periodic rebalancing produces significantly higher Sharpe ratios compared to buy-and-hold if the stocks portfolio weight accounts for at least 30%. If the relative proportion of stocks falls below 30%, no statistical inferences can be drawn because the difference between rebalancing and buy-and-hold is lost in estimation error. In contrast to the United States, the optimal asset allocation comprises of 20% stocks and 80% government bonds. Moreover, adopting a buy-and-hold strategy, the average Sharpe ratio of the optimal stock-bond portfolio with a value of 0,478 is substantially lower compared to the United States.

Panel E of Figure I shows the average Sharpe ratios of both quarterly periodic rebalancing and buy-and-hold across different asset allocations for the financial market of Germany, while Panel F of Figure I depicts the corresponding 10%-quantiles. Our history-based simulation results reconfirm the finding that rebalancing leads to a better risk-adjusted performance in terms of average Sharpe ratios across different asset allocations. We even observe statistical significance if the stock allocation exceeds 0%. Moreover, as hypothesized by the discussion of the descriptive statistics in Table I, the optimal stock-bond portfolio of the German financial market consists of a higher proportion of government bonds in comparison to the optimal portfolio mix of the United States and the United Kingdom.

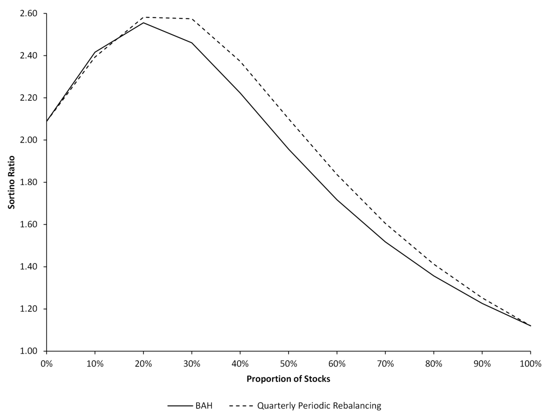
4.6.2 Sortino Ratio

In contrast to the Sortino ratio, the Sharpe ratio considers all deviations from the mean return – both negative and positive. However, as positive deviations constitute an opportunity to generate an additional return on the invested capital, they are not expected to be perceived as risk by investors. Defined by Sortino and Price (1994), the Sortino ratio should better reflect investors' risk perception by representing a risk-adjusted performance measure that only takes negative deviations into account:

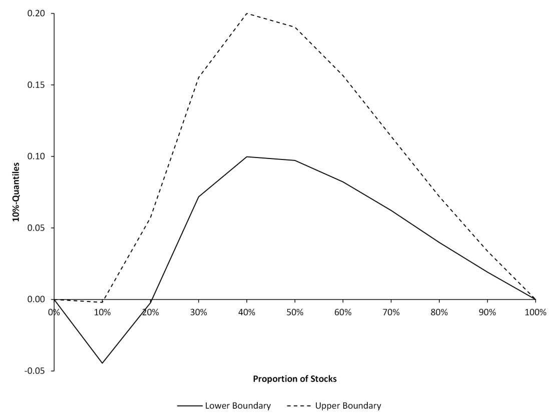
$$S_i(\tau) = \frac{\bar{r}_i - \tau}{\sqrt{\int_{-\infty}^{\tau} (\tau - r_i)^2 f(r_i) dr_i}}, \tag{7}$$

where $\bar{r}_i$ is the average return of strategy i, $f(\cdot)$ the corresponding probability density function, and τ the threshold return required by the investor. Illustrating the average Sortino ratio of quarterly periodic rebalancing and buy-and-hold for the United States, Panel A and Panel B of Figure II reconfirm our findings of the results stated above.

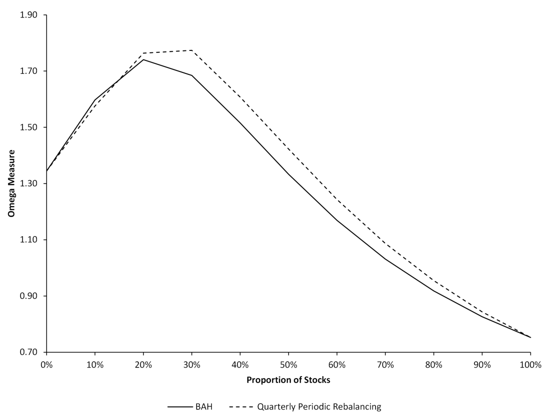
Figure II – Sortino Ratio and Omega Measure Across Different Asset Allocations



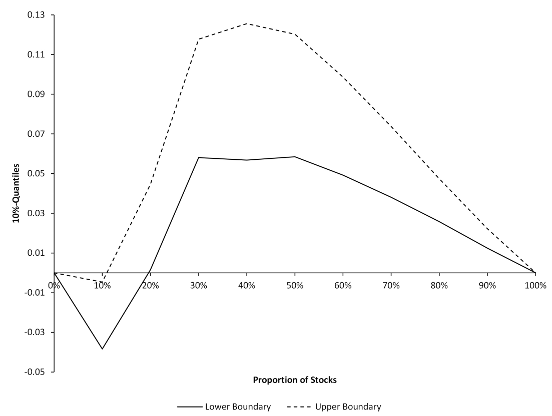
(A) Average Sortino Ratio (United States)



(B) 10%-Quantiles of the Average Sortino Ratio



(C) Average Omega Measure (United States)



(D) 10%-Quantiles of the Average Omega Measure

4.6.3 Omega Measure

Focusing on a two-asset class portfolio consisting of stocks and government bonds, portfolio returns are a linear function of stock returns if no transaction takes place during the investment period. Thus, buy-and-hold represents a linear investment strategy. In contrast, rebalancing constitutes a concave investment strategy as illustrated by Perold and Sharpe (1988). On the one hand, the frequent reallocation to the weaker performing asset reduces the upside potential in upswing markets, thereby leading to an increase of the portfolio return at a declining rate. On the other hand, a regular readjustment to the target weights also reduces the downside protection in downswing markets, which causes a decline of the portfolio return at an increasing rate. However, Ingersoll et al. (2007) document that due to construction, investment strategies with a concave payoff lead to higher Sharpe ratios and Sortino ratios in comparison to buy-and-hold. Although this result does not contradict the finding that rebalancing outperforms buy-and-hold, it points out that the possibility of portfolio performance manipulation as stated by Ingersoll et al. (2007) must not be neglected.

For this reason, we further apply the Omega measure as a robustness check. Developed by Kaplan and Knowles (2004), the Omega measure represents the ratio of gains to losses relative to a predetermined threshold return:

$$\Omega_i(\tau) = \frac{\int_{\tau}^{\infty}(1 - F(r_i))dr_i}{\int_{-\infty}^{\tau} F(r_i)dr_i}, \tag{8}$$

where τ denotes the return of strategy i, $F(\cdot)$ the cumulative density function, and τ the predetermined threshold return. In contrast to the Sharpe ratio and the Sortino ratio, the Omega measure considers the entire return distribution of the underlying investment strategy. As all moments are taken into account, a portfolio performance manipulation can be considered extremely difficult. Therefore, the Omega measure contributes to careful evaluation of risk-adjusted portfolio performance.

By way of example, Panel C of Figure II plots the average Omega measures of quarterly periodic rebalancing and buy-and-hold against the underlying portfolio mix. Panel D of Figure II shows the corresponding 10%-quantiles. All in all, Panel C and Panel D of Figure II substantiate our prior findings that rebalancing significantly boosts risk-adjusted portfolio performance if stocks' proportion exceeds a certain threshold. This proportion is subject to the country, the period under investigation (results not reported), and the risk-adjusted performance measure. In the case of Omega measures, it amounts to 20% for the United States and the United Kingdom and 0% for Germany. In results not reported, we observe qualitatively similar results for the financial markets of the United Kingdom and Germany.

Overall, our analysis points out that cross-country differences have a substantial impact on the level of risk-adjusted performance. Moreover, the period under investigation also affects the level as well as the optimal asset allocation (results not reported). However, provided that the stock allocation exceeds a certain threshold which is subject to the underlying country, our history-based simulation provides strong evidence that rebalancing leads to a value added for institutional investors across all countries.

4.7 Conclusion

This paper compares the impact between rebalancing and buy-and-hold on the risk-adjusted performance of stock-bond portfolios across different asset allocations by reporting statistical significance levels. In order to simulate realistic market conditions and derive valuable recommendations for investment practice, we apply the stationary bootstrap of Politis and Romano (1994) which enables us to preserve most of the time series characteristics of financial time series (e.g. short-term momentum, fat tails, and empirical distribution). Modifying the stocks' portfolio weight in 10% steps from 0% to 100% stocks, our analysis comprises eleven different stock-bond portfolios for each of the three countries under investigation.

Our results provide strong evidence that rebalancing leads to a superior risk-adjusted performance across almost all asset allocations if the stocks' portfolio weight exceeds a certain threshold. However, the exact threshold ranges between 0% and 30% and is subject to both the country as well as to the applied risk-adjusted performance measure. Despite substantial cross-country differences, our analysis documents that rebalancing provides a value added to institutional investors of all three countries across different asset allocations. However, the optimal asset allocation strongly depends on both the country and the period under investigation.

References

Adcock, C.J., N. Areal, M.J.R. Armada, M.C. Cortez, B. Oliveira, and F. Silva (2012). "Conditions under Which Portfolio Performance Measures Are Monotonic Functions of the Sharpe Ratio." Working Paper, University of Sheffield Management School, Sheffield, 2012.

Arnott, R.D. and R.M. Lovell (1993). "Rebalancing: Why? When? How Often?" In: *Journal of Investing* 2, pp. 5–10.

Black, F. and R. Litterman (1992). "Global Portfolio Optimization." In: *Financial Analysts Journal* 48, pp. 28–43.

Brinson, G.P., L.R. Hood, and G.L. Beebower (1986). "Determinants of Portfolio Performance." In: *Financial Analyst Journal* 42, pp. 39–44.

Brinson, G.P., B.D. Singer, and G.L. Beebower (1991). "Determinants of Portfolio Performance II: An Update." In: *Financial Analyst Journal* 47, pp. 40–48.

Brock, W., J. Lakonishok, and B. LeBaron (1992). "Simple Technical Trading Rules and the Stochastic Properties of Stock Returns." In: *Journal of Finance* 47, pp. 1731–1764.

DeMiguel, V., L. Garlappi, F.J. Nogales, and R. Uppal (2009a). "A Generalized Approach to Portfolio Optimization: Improving Performance by Constraining Portfolio Norms." In: *Management Science* 55, pp. 798–812.

DeMiguel, V., L. Garlappi, and R. Uppal (2009b). "Optimal Versus Naive Diversification: How Inefficient Is the 1/N Portfolio Strategy?" In: *Review of Financial Studies* 22, pp. 1915–1953.

Dichtl, H., W. Drobetz, and M. Wambach (2013). "Testing Rebalancing Strategies for Stock-Bond Portfolios: What Is the Optimal Rebalancing Strategy?" Working Paper, University of Hamburg, Hamburg, 2013.

Efron, B. and R.J. Tibshirani (1998). *An Introduction to the Bootstrap*. Chapman & Hall/CRC.

Eling, M. (2008). "Does the Measure Matter in the Mutual Fund Industry?" In: *Financial Analyst Journal* 64, pp. 54–66.

Fabozzi, F.J., P.N. Kolm, D.A. Pachamanova, and S.M. Focardi (2007). "Robust Portfolio Optimization." In: *Journal of Portfolio Management* 33, pp. 40–48.

Harjoto, M.A. and F.J. Jones (2006). "Rebalancing Strategies for Stocks and Bonds Asset Allocation." In: *Journal of Wealth Management* 9, pp. 37–44.

Ibbotson, R.G. and P.D. Kaplan (2000). "Does Asset Allocation Policy Explain 40, 90, or 100 Percent of Performance?" In: *Financial Analyst Journal* 56, pp. 26–33.

Ingersoll, J., M. Spiegel, W. Goetzmann, and I. Welch (2007). "Portfolio Performance Manipulation and Manipulation-Proof Performance Measures." In: *Review of Financial Studies* 20, pp. 1503–1546.

Jaconetti, C.M., F.M. Kinniry, and Y. Zilbering (2010). *Best Practices for Portfolio Rebalancing*. Tech. rep. Vanguard.

Kaplan, P.D. and J.A. Knowles (2004). "Kappa: A Generalized Downside Risk-Adjusted Performance Measure." In: *Journal of Performance Measurement* 8, pp. 42–54.

Kritzman, M. (2006). "Are Optimizers Error Maximizers?" In: *Journal of Portfolio Management* 32, pp. 66–69.

Ledoit, O. and M. Wolf (2008). "Robust Performance Hypothesis Testing with the Sharpe Ratio." In: *Journal of Empirical Finance* 15, pp. 850–859.

– (2011). "Robust Performance Hypothesis Testing with Variance." In: *Wilmott Magazine* 55, pp. 86–89.

Leland, H. E. (1999). "Optimal Portfolio Management with Transactions Costs and Capital Gains Taxes." Research Program in Finance Working Paper RPF-290, University of California, Berkeley, 1999.

Markowitz, H. (1952). “Portfolio Selection.” In: *Journal of Finance* 7, pp. 77–91.

Michaud, R.O. (1989). “The Markowitz Optimization Enigma: Is Optimized Optimal?” In: *Financial Analysts Journal* 45, pp. 31–42.

Norwegian Ministry of Finance (2012). *Report No. 17 to the Storting: The Management of the Government Pension Fund in 2011*. URL: http://www.regjeringen.no/pages/37635499/stm_gpf.pdf.

Patton, A., D.N. Politis, and H. White (2009). “Correction to “Automatic Block-Length Selection for the Dependent Bootstrap” by D. Politis and H. White.” In: *Econometric Reviews* 28, pp. 372–375.

Perold, A.F. and W.F. Sharpe (1988). “Dynamic Strategies for Asset Allocation.” In: *Financial Analysts Journal* 44, pp. 16–27.

Politis, D.N. and J.P. Romano (1994). “The Stationary Bootstrap.” In: *Journal of the American Statistical Association* 89, pp. 1303–1313.

Politis, D.N. and H. White (2004). “Automatic Block-Length Selection for the Dependent Bootstrap.” In: *Econometric Reviews* 23, pp. 53–70.

Romano, J.P. and M. Wolf (2006). “Improved Nonparametric Confidence Intervals in Time Series Regressions.” In: *Journal of Nonparametric Statistics* 18, pp. 199–214.

Sortino, F.A. and L.N. Price (1994). “Performance Measurement in a Downside Risk Framework.” In: *Journal of Investing* 3, pp. 59–64.

Tokat, Y. and N. Wicas (2007). “Portfolio Rebalancing in Theory and Practice.” In: *Journal of Investing* 16, pp. 52–59.

Tsai, C.S.-Y. (2001). “Rebalancing Diversified Portfolios of Various Risk Profiles.” In: *Journal of Financial Planning* 14, pp. 104–110.

Xiong, J.X., R.G. Ibbotson, T.M. Idzorek, and P. Chen (2010). “The Equal Importance of Asset Allocation and Active Management.” In: *Financial Analysts Journal* 66, pp. 22–30.

Chapter 5

A Bootstrap-Based Comparison of Portfolio Insurance Strategies

with H. Dichtl and W. Drobetz

Abstract

This study presents a systematic comparison of portfolio insurance strategies. In order to test for statistical significance of the differences in Omega ratios between different pairs of portfolio insurance strategies, we use a bootstrap-based hypothesis test proposed by Politis and Romano (1994). Our comparison of the different strategies considers the following distinguishing characteristics: static vs. dynamic; initial wealth vs. cumulated wealth protection; model-based vs. model-free; and strong floor compliance vs. probabilistic floor compliance. Overall, the classical portfolio insurance strategies synthetic put and CPPI provide superior downside protection compared to a simple stop-loss trading rule, which also results in significantly higher Omega ratios. Analyzing more recently developed strategies, neither the TIPP strategy (as an 'improved' CPPI strategy) nor the dynamic VaR-strategy are able to provide significant improvements over the classical portfolio insurance strategies. Most importantly, when parameter estimation risk is incorporated into the analysis, our comparison of the dynamic VaR-strategy with the CPPI strategy and the synthetic put strategy shows no value added of this novel approach.

5.1 Introduction

Driven by their risk and return preferences or by regulatory requirements, many investors are highly interested in preserving their invested capital. Therefore, portfolio insurance is of utmost importance for the asset management industry. The major objective of any portfolio insurance strategy is to limit the downside risk of a risky asset (or a portfolio of risky assets), while simultaneously maintaining most of the upside return potential. Institutional investors often use portfolio insurance strategies in tailor-made solutions to protect their portfolios against large losses, whereas many private investors invest their capital in mutual funds that are endowed with a capital protection guarantee. The literature offers a number of different methods and strategies to secure risky assets against large losses, such as the stop-loss strategy, the synthetic put strategy, or the constant proportion portfolio insurance (CPPI) strategy. Therefore, once an investor has decided to use a protection strategy within his asset allocation framework, it may be of critical importance to implement the 'best' strategy which fits the investor's preferences as good as possible.

The introduction of synthetic put strategies by Rubinstein and Leland (1981) has laid a foundation for numerous studies that contribute to explain as to why portfolio insurance is so popular in investment practice. Prior studies investigate the optimality of portfolio insurance strategies in a standard expected utility maximization set-up (Leland (1980), Brennan and Solanki (1981), Benninga and Blume (1985), Dybvig

(1988), Black and Litterman (1992)).[1] A consensus in this strand of the literature is that portfolio insurance strategies are only utility maximizing under very specific assumptions with respect to investor preferences. In a more recent strand of the literature, Annaert et al. (2009) document that portfolio insurance strategies outperform a buy-and-hold strategy in terms of downside protection, but provide lower excess returns. A comparison based on stochastic dominance criteria reveals that no dominance relations can be identified between portfolio insurance strategies and buy-and-hold.[2] Given that their lower return potential is sufficiently compensated by lower risk, Annaert et al. (2009) conclude that portfolio insurance strategies are valuable alternatives at least for some investors. Dichtl and Drobetz (2011) analyze portfolio insurance strategies within the framework of the cumulative prospect theory. Their simulation results indicate that loss aversion and probability weighting contribute to make most portfolio insurance strategies a preferred investment strategy for a prospect theory investor. Dierkes et al. (2010) similarly document that portfolio insurance is attractive for an investor having preferences described by the cumulative prospect theory.

Given the strategic decision to implement a portfolio insurance strategy, the question of interest is which particular strategy should be adopted. In fact, several prior studies evaluate and compare different pairs of portfolio insurance strategies in terms of their risk and return characteristics. For example, Benninga (1990) compares the stop-loss rule, the CPPI strategy, and the synthetic put strategy by using a Monte Carlo simulation approach. He reports that the simple stop-loss rule dominates both the CPPI strategy and the synthetic put strategy in terms of their terminal wealth and Sharpe ratio.[3] Cesari and Cremonini's (2003) simulation results indicate that the relative performance of portfolio insurance strategies depends on the market environment. For example, using different performance measures (including alternative risk measures, such as the downside deviation and the Sortino ratio), the CPPI strategy dominates all other strategies in bear and sideway markets.[4]

Herold et al. (2005) examine a value-at-risk-based (VaR) dynamic investment strategy,

---

[1] Brennan and Schwartz (1989), Basak (1995), and Grossman and Zhou (1996) analyze the equilibrium implications of portfolio insurance in an expected utility framework.

[2] Annaert et al. (2009) implement block-bootstrap simulations from the empirical distributions to incorporate heavy tails and volatility clustering. Bertrand and Prigent (2005) and Zagst and Kraus (2011) also evaluate portfolio insurance performance using stochastic dominance. For example, Zagst and Kraus (2011) derive parameter conditions implying the second- and third-order stochastic dominance of the CPPI strategy against the protective put strategy. However, they assume that the underlying risky asset follows a geometric Brownian motion.

[3] Do (2002) compares the CPPI strategy and the synthetic put strategy using Australian market data. While neither strategy can be justified based on a loss minimization or a gain participation point of view, he nevertheless reports that the CPPI strategy dominates in terms of floor protection and the costs of insurance.

[4] Bookstaber and Langsam (2000) also provide a detailed analysis of the path dependency of portfolio insurance strategies.

where the protection floor is maintained with a pre-specified probability (probabilistic floor compliance). They compare this strategy with the CPPI strategy and document a higher return of the dynamic VaR-strategy in the context of fixed-income investments.[5] Jiang et al. (2009) also compare a variant of a dynamic VaR-based portfolio insurance strategy with the CPPI strategy and buy-and-hold. Their results indicate that the dynamic VaR-based strategy outperforms the CPPI strategy in terms of downside protection and return potential. A caveat is that Jiang et al. (2009) approximate the expected return and volatility using their realized values during the insurance period, and thus their analysis neglects the estimation risk inherent in a dynamic VaR-strategy. Most recently, Bertrand and Prigent (2011) suggest that the Omega measure is the most adequate performance measure because it is able to capture the entire empirical return distribution of portfolio insurance strategies. Using a block-bootstrap simulation approach, they compare the CPPI strategy with the synthetic put strategy and report a dominance of the CPPI strategy in terms of the Omega ratio.

This study complements and expands these earlier studies which evaluate and compare different portfolio insurance strategies. We not only analyze different pairs of portfolio insurance strategies in terms of their protection quality and return potential, but rather, provide a comprehensive statistical analysis of five protection strategies: the traditional stop-loss strategy, the synthetic put strategy, the CPPI strategy, the time invariant portfolio protection (TIPP) strategy, and the VaR-based protection strategy. Using a bootstrap simulation approach based on historical data as well as scenario-based Monte Carlo simulation, we investigate the choice between the following characteristics of different portfolio insurance strategies: (1) static versus dynamic protection; (2) initial wealth versus accumulated wealth protection; (3) model-based versus model-free protection; and (4) strong floor compliance versus probabilistic floor compliance. In each case, we compare two adequate portfolio insurance strategies by using Shadwick and Keating's (2002) Omega ratio as the appropriate risk-adjusted performance measure (Bertrand and Prigent (2011)). By implementing a bootstrap-based hypothesis test suggested by Politis and Romano (1994), which does not require specific assumptions about the distributional properties of the Omega ratios themselves, we are able to provide statistical inference. In addition, for those strategies which require parameter estimates (i.e., the synthetic put strategy and the dynamic VaR-strategy), we incorporate estimation risk.

Our results indicate that the 'classical' portfolio insurance strategies CPPI and synthetic put dominate all other strategies under realistic investment conditions. Both strategies are superior to the simple stop-loss trading rule in terms of their protection quality and overall risk-adjusted performance measured by the Omega ratio.

[5] Given the comparatively low downside potential of fixed-income portfolios, it may be difficult to assess the protection quality of the dynamic VaR-strategy in Jiang et al.'s (2009) framework.

Furthermore, recently developed portfolio insurance strategies also seem to be at a disadvantage against the more standard variants. Specifically, while the TIPP strategy (as an 'improved' CPPI strategy) exhibits a very limited return potential, the usefulness of the dynamic VaR-strategy strongly depends on the forecasting quality for both the equity risk premium and the stock market volatility. The remainder of this paper is structured as follows: Section 5.2 provides an overview and a brief discussion of the implemented portfolio insurance strategies. While Section 5.3 describes our historical simulation methodology and our test design for statistical significance, Section 5.4 presents the historical simulation results. As a robustness test, Section 5.5 documents the methodology and results of Monte Carlo simulations. Section 5.6 concludes and emphasizes the implications for the investment practice.

5.2 An Overview of Portfolio Insurance Strategies

A variety of portfolio insurance strategies have been suggested in the prior literature, such as the stop-loss trading rule, the use of derivatives (e.g., a protective put strategy), and dynamic trading strategies (e.g., a synthetic put strategy or the CPPI strategy). In contrast to the stop-loss trading rule and dynamic trading strategies, a protective put strategy requires a liquid put option with the desired strike price and the desired time to maturity (Figlewski et al. (1993)). Therefore, we omit option-based portfolio insurance strategies and focus on the simple stop-loss trading rule and dynamic trading strategies.[6]

5.2.1 Stop-Loss Portfolio Insurance Strategy

The simplest way to protect a risky portfolio against losses is the stop-loss portfolio insurance strategy (Bird et al. (1988)). An investor invests his total wealth (W_0) in the risky asset. This position is maintained as long as the market value of the risky position (W_t) exceeds the net present value (NPV) of the floor (F_T), which is the minimum acceptable portfolio value at the end of the investment horizon T:

$$W_t > NPV(F_T). \tag{1}$$

If the portfolio value reaches or drops below the discounted floor, i.e., if $W_t = NPV(F_T)$, all risky portfolio holdings are sold and invested in the risk-free asset. This position is held until the end of the investment horizon. As long as the interim portfolio value does not drop below $NPV(F_T)$, an investor's final wealth will never be lower than F_T. Taken together, the stop-loss portfolio insurance strategy is easy to implement, it

[6] The following description of portfolio insurance strategies is based on Dichtl and Drobetz (2011).

does not depend on any specific assumptions, and it also does not require estimating any model parameters. However, the investor no longer participates from any upward market movement once the portfolio has been shifted in the risk-free asset.

5.2.2 Synthetic Put Portfolio Insurance Strategy

A second portfolio insurance strategy is Rubinstein's (1981) synthetic put strategy, which uses the Black and Scholes (1973) option pricing formula to create a continuously adjusted synthetic European put option on the risky asset. Combining the purchase of the risky asset with the purchase of a put on this asset (stock) is equivalent to purchasing a continuously-adjusted portfolio which is a combination of the risky asset and the risk-free asset (cash). The value of a portfolio that consists of a stock S plus a put P can be calculated as:

$$\begin{aligned} S + P &= S - S \cdot \mathcal{N}(-d_1) + K \cdot e^{-rT} \mathcal{N}(-d_2) \\ &= S \cdot [1 - \mathcal{N}(-d_1)] + K \cdot e^{-rT} \mathcal{N}(-d_2) \\ &= S \cdot \mathcal{N}(d_1) + K \cdot e^{-r} \mathcal{N}(-d_2), \end{aligned} \tag{2}$$

where K is the strike price, r the risk-free rate, and T the time to maturity. $\mathcal{N}(\cdot)$ is the standard normal cumulative distribution function with d_1 and d_2 defined as:

$$d_1 = \frac{ln(S/K) + (r + 0.5\sigma^2)T}{\sigma\sqrt{T}}, \tag{3a}$$

$$d_2 = d_1 - \sigma\sqrt{T}, \tag{3b}$$

where σ denotes the standard deviation of risky asset returns. In order to calculate the investment in the risky asset of the replicating portfolio, the delta of the portfolio in equation (2) is:

$$\frac{\partial(S + P)}{\partial S} = \mathcal{N}(d_1). \tag{4}$$

Multiplying the delta with the price of the risky asset S and dividing by the value of the portfolio in equation (2), the percentage allocations in the risky asset and the risk-free asset are:

$$w_{\text{risky}} = \frac{S \cdot \mathcal{N}(d_1)}{S \cdot \mathcal{N}(d_1) + K \cdot e^{-rT} \mathcal{N}(-d_2)}, \tag{5a}$$

$$w_{\text{risk-free}} = 1 - w_{\text{risky}}. \tag{5b}$$

This strategy requires increasing (decreasing) the proportion of the risky asset in the portfolio if the price of the risky asset increases (decreases). In order to maintain a desired protection level, the strike price K must be set such that the following relationship holds:

$$K = \frac{F_T}{W_0} \cdot (S + P(K)), \tag{6}$$

where the ratio F_T/W_0 is the percentage floor. The solution of equation (6) must be determined iteratively, because the value of the put option $P(K)$ depends on the strike price itself. Given that the put itself is costly, the exercise price K of the put will be higher than the floor.

The portfolio must be readjusted on a continuous basis in order to maintain the desired protection level. However, transaction costs will incur with each portfolio adjustment, and thus a higher adjustment frequency leads to higher transaction costs. In order to incorporate this transaction costs effect, Leland (1985) as well as Boyle and Vorst (1992) suggest using the synthetic put portfolio insurance strategy with a modified volatility estimator:

$$\sigma_{\text{Leland}} = \sigma\sqrt{1 + \sqrt{\frac{2}{\Pi}} \cdot \frac{k}{\sigma\sqrt{\Delta t}}}, \tag{7a}$$

$$\sigma_{\text{Boyle/Vorst}} = \sigma\sqrt{1 + 2 \cdot \frac{k}{\sigma\sqrt{\Delta t}}}, \tag{7b}$$

where k captures the round-trip transaction costs, and Δt denotes the length of the readjustment period. With these volatility adjustments, both the readjustment frequency and the corresponding transaction costs are taken into account.

5.2.3 Constant Proportion Portfolio Insurance Strategy

In contrast to a synthetic put strategy, the constant proportion portfolio insurance strategy (CPPI) is a dynamic protection strategy that is not based on option pricing theory (Black and Jones (1987)). Starting point is an investor's risk capital at time t (the 'cushion'). The current cushion C_t represents the difference between current wealth at time t, labeled W_t, and the discounted floor $NPV(F_T)$:

$$C_t = W_t - NPV(F_T). \tag{8}$$

The exposure into the risky asset at time t, denoted as e_t, is calculated by multiplying the cushion C_t with the multiplier m:

$$e_t = m \cdot C_t. \tag{9}$$

The multiplier m could be set to any value, but it has a strong economic implication. Specifically, the inverse of the multiplier ($1/m$) represents the maximum sudden loss in the risky asset that may occur such that the cushion is not fully depleted and the portfolio value does not fall below the discounted floor.[7] Although maintenance of the floor will be controlled on an intraday basis, the risk still exists over night when the portfolio manager cannot react immediately to extreme market losses ('overnight risk' or 'gap risk'). In spite of this small (unlikely) residual risk, which can partly be controlled by choosing a small multiplier m, the CPPI strategy can be classified as an absolute protection strategy with a strictly lower limit on the portfolio value. In commercial applications, the CPPI strategy is usually implemented such that the exposure of the risky asset varies between 0% and 100%. This implies that short sales and leverage are excluded (Benninga (1990), Do (2002), Annaert et al. (2009)):

$$e_t = \max[\min(m \cdot C_t, W_t), 0]. \tag{10}$$

The CPPI strategy can be implemented in our simulation analysis as dictated in equation (10) by shifting between stocks and bills.

5.2.4 Time Invariant Portfolio Protection Strategy

Estep and Kritzman (1988) argue that investors will not only be interested in the protection of their initial wealth, but also in locking in all interim capital gains. They suggest a modification of the standard CPPI strategy, which they call the 'time invariant portfolio protection' (TIPP) strategy. While the CPPI strategy operates with a fixed floor (i.e., the initial wealth multiplied by the percentage floor), the floor of the TIPP strategy is ratcheted up if the value of the portfolio increases. After choosing the floor and the multiplier, this strategy requires the following steps (Estep and Kritzman (1988)):

1. Calculation of the actual portfolio value (stock plus cash).
2. Multiplication of this portfolio value by the floor percentage.

---

[7] For example, with a multiplier of $m = 5$, the risky asset can lose 20% ($1/5 = 0.20$) without violating the floor. When a sudden loss of over 20% occurs, the value of the portfolio falls below the promised minimum value ('gambler's ruin'). In commercial applications it is necessary to continuously control the optimal exposure in the risky asset, and thus portfolio shifts would have to be executed immediately. In most instances, however, an appropriate trading filter is used.

3. If the result in step 2 is greater than the previous floor, it becomes the new floor; otherwise the old floor is kept.

4. Application of the CPPI strategy as dictated by equations (8)-(10).

As in the CPPI strategy, the TIPP strategy transfers all risky asset holdings in an irreversible manner to the risk-free asset once the floor has been reached. Therefore, the TIPP strategy cannot participate from subsequent upward market movements. Given the continuous 'ratcheting up' of the floor to the highest portfolio value, the likelihood that the portfolio value reaches or violates the prevailing floor increases, potentially implying adverse return effects because the TIPP strategy may more often end up fully invested in the risk-free asset (Choie and Seff (1989), Dichtl and Drobetz (2011)).

5.2.5 Dynamic Value-at-Risk Portfolio Insurance Strategy

The target of the dynamic VaR-based strategy is to control the exposure of a risky asset such that a specified value at risk is not violated (Jiang et al. (2009)). The value-at-risk (VaR) of any portfolio P defines the maximum loss that will not be exceeded based on a defined confidence level within a given time period:

$$VaR = T \cdot \mu_P + \tau_{1-\alpha} \cdot \sqrt{T} \cdot \sigma_P, \tag{11}$$

where T is the time horizon and $\tau_{1-\alpha}$ the $(1-\alpha)$-quantile of the standard normal distribution. μ_P and σ_P denote the expected portfolio return and volatility, respectively. Consider a risk-free position (with return r_f) and a stock (market) position (with expected return μ_S and volatility σ_S), where x represents the percentage stock market allocation. Given that the risk-free asset has zero volatility and is uncorrelated with stock returns, the value-at-risk is derived as follows:

$$VaR = T \cdot [x \cdot (\mu_S - r_f) + r_f] + \tau_{1-\alpha} \cdot \sqrt{T} \cdot x \cdot \sigma_S. \tag{12}$$

Within this VaR-based asset allocation, the question of interest is how to allocate portfolio wealth between stocks and cash in order to ensure that a given VaR-limit will not be violated. More specifically, the value-at-risk in equation (12) is set to a pre-specified value, and the expression is solved for the stock market allocation x:

$$x = \frac{VaR - T \cdot r_f}{T \cdot (\mu_S - r_f) + \tau_{1-\alpha} \cdot \sqrt{T} \cdot \sigma_S}. \tag{13}$$

Various parameters must be specified in equation (13): the confidence level α, the risk-free rate r_f, the expected stock market volatility σ_S, and the equity risk premium $(\mu_S - r_f)$. Both the future stock market volatility and the expected risk premium are ex ante unobservable and need to be estimated. Instead of applying equation (13) only once at the beginning of the insurance period, an investor needs to continuously readjust the stock-cash allocation. Thus, the VaR-strategy constitutes a dynamic trading strategy comparable to the CPPI or the synthetic put strategy.[8] Commercial applications avoid leverage and short sales, and thus the percentage stock market allocation x in equation (13) varies between 0% and 100% in our simulation analysis.

5.2.6 Distinguishing Characteristics of Portfolio Insurance Strategies

The main properties of the different portfolio insurance strategies discussed in this section are summarized in Table I. First of all, portfolio insurance strategies can be classified into static and dynamic strategies. While static strategies maintain their initial asset allocation or change it only once during the entire insurance period, dynamic protection strategies adjust their allocation continuously on the basis of market movements. Therefore, the stop-loss trading rule is classified as a static strategy, while all other trading rules are dynamic strategies. A second distinguishing feature is the protection level. While the TIPP strategy (as a modified CPPI strategy) locks in the cumulated wealth, all other strategies set their protection target at the initial portfolio wealth. A third criterion for differentiation is the type of insurance provided. In contrast to the strategies stop-loss, synthetic put, CPPI, and TIPP, a dynamic VaR-strategy maintains the floor only with a pre-specified probability. We refer to this feature as 'probabilistic floor compliance' as opposed to 'strong floor compliance' of all other portfolio insurance strategies in Table I. The fourth criterion for differentiation relates to the model parameters. While some strategies require one or two model parameters (stop-loss, CPPI, and TIPP), other strategies depend on more model parameters (synthetic put and dynamic VaR-strategy). Furthermore, both the synthetic put and the dynamic VaR-strategy include forward looking model parameters (Bird et al. (1990), Zhu and Kavee (1988), Jiang et al. (2009)), and therefore the precision of these parameter estimates will strongly impact the quality of loss protection (Rendleman and OBrien (1990)). Finally, the various strategies can be differentiated in terms of their model assumptions. While stop-loss, CPPI, and TIPP do not depend on any model assumptions, the synthetic put strategy and the dynamic VaR-strategy are based on normally distributed returns.

[8] It is important to consider cumulated portfolio performance when implementing the dynamic VaR-strategy. For example, if the target-VaR over a one-year investment horizon is 0% and the portfolio wealth has increased from 100 to 110 after one month, the VaR-budget is −9.09%, and thus a loss of up to −9.09% does not violate the original VaR-budget. However, if portfolio wealth has decreased from 100 to 90, the new VaR-budget becomes +11.11%. The portfolio value must increase by at least 11.11% in order to maintain the original VaR-budget of 0%.

Table I – Main Properties of Portfolio Insurance Strategies

	Stop-Loss	Synthetic Put	CPPI	TIPP	Dynamic VaR
Insurance strategy	Static	Dynamic	Dynamic	Dynamic	Dynamic
Protection target	Initial wealth	Initial wealth	Initial wealth	Cumulated wealth	Initial wealth
Type of insurance	Strong floor compliance	Strong floor compliance	Strong floor compliance	Strong floor compliance	Probabilistic floor compliance
Model parameters					
Fixed	• Floor	• Stock price (S) • Strike price (K) • Risk-free rate (r_f) • Time-to-maturity (T)	• Floor • Multiplier (m)	• Floor • Multiplier (m)	• VaR (Floor) • Risk-free rate (r_f • Confidence level (α) • Equity risk premium ($\mu_S - r_f$)
Estimators	• Volatility (σ_S)				• Volatility (σ_S)
Model assumptions	No assumptions	Normal distribution	No assumptions	No assumptions	Normal distribution

Taken together, our systematic comparison of the different portfolio insurance strategies is structured along the following distinguishing characteristics:

Static vs. dynamic portfolio insurance: In a first step of our simulation analysis, we compare the simple stop-loss trading rule with the CPPI strategy (comparison 1). Apart from the difference between static and dynamic protection, the two strategies are equal along all other dimensions. Both strategies focus on the protection of initial wealth, provide strong floor compliance, are free of any model assumptions, and do not require the estimation of model parameters.

Initial wealth vs. cumulated wealth protection: In a second step, we compare the CPPI strategy with the related TIPP strategy (comparison 2). The latter TIPP strategy is a modified CPPI strategy, where the only difference relates to the level of wealth protection. Both strategies are dynamic, guarantee a strong floor compliance, are independent from any model assumptions, and do not require the forecast of any model parameters.

Model-based vs. model-free portfolio insurance: In order to assess model-based against model-free protection strategies, we compare the synthetic put with the CPPI strategy in a third step (comparison 3). While both strategies are dynamic, their main difference relates to the model assumptions. In contrast to the CPPI strategy, the synthetic put strategy is derived from the Black and Scholes (1973) option pricing model and depends on its underlying assumptions. Moreover, the synthetic put trading rule requires estimating the stock market volatility to be used in the option pricing formula.

Strong floor compliance vs. probabilistic floor compliance: In the final step, we examine the dynamic VaR-strategy in comparison to the synthetic put strategy (comparison 4a). Both strategies are dynamic, assume normally distributed stock returns, and require an estimation of unobservable strategy parameters. Furthermore, we compare the dynamic VaR-strategy with the CPPI strategy (comparison 4b). While both strategies are dynamic, the CPPI strategy is independent of any model assumptions and parameter estimates. Taken together, comparing these two pairs of strategies highlights the difference between strong floor compliance (synthetic put and CPPI strategy) and probabilistic floor compliance (dynamic VaR-strategy).

5.3 Historical Simulation Methodology

5.3.1 Data and Design of the Empirical Analysis

Our data consists of daily returns for the German stock market index DAX and money market rates from the German Bundesbank (average values of the 1-month Frankfurt interbank rate; middle rate) over the sample period from January 1981 to December 2011. As documented in Poterba and Summers (1988), among others, stock market returns exhibit short-term autocorrelation and long-term mean reversion. Moreover, stock market returns deviate from the normal distribution with constant volatility (homoscedasticity); they tend to be heteroscedastic and left-skewed, and they exhibit fat tails (Annaert et al. (2009), Bertrand and Prigent (2011)). As shown in Dichtl and Drobetz (2011), similar statistical properties are also observable for daily DAX return series. In order to fully incorporate the time series characteristics, we follow Annaert et al. (2009) and Bertrand and Prigent (2011) and implement a block-bootstrap by drawing 250 subsequently following pairs of daily stock and money market returns.[9] Selecting $N = 5,000$ coherent data blocks with replacement and applying the different portfolio insurance strategies in each draw provides us with $5,000$ annual returns for each strategy.[10] These $N = 5,000$ yearly returns are the basis for our performance measures (see Section 5.3.2) as well as our hypothesis tests (see Section 5.3.3).

The equity risk premium and the volatility for each one-year simulation path of the dynamic VaR-strategy are estimated from the prior 250 daily returns. The return volatility for the synthetic put strategy (again estimated from the prior 250 daily returns) is modified according to Boyle and Vorst (1992), as shown in equation (7b).[11] We implement all strategies without short sales and leverage and take round-trip transaction costs of 10 basis points into account (Herold et al. (2007)). In order to avoid portfolio shifts that are triggered by trendless market movements, we implement the dynamic VaR-strategy, the CPPI strategy, the TIPP strategy, and the synthetic put strategy with a trading filter. Portfolio shifts are only executed when the stock market moves by more than 2% (Do and Faff (2004)). Moreover, in the basic model specification, we implement both the CPPI strategy and the TIPP strategy with a multiplier of $m = 5$, which is commonly used in commercial applications (Herold et al. (2007)). The confidence level in the dynamic VaR-strategy is $a = 90\%$. We implement all strategies based on a protection level of 100% (i.e., with a full capital guarantee) and compare their performance with an unprotected buy-and-hold stock market investment.

5.3.2 Performance Measures

In order to compare the different portfolio insurance strategies as described in Section 5.2.1-5, we investigate their protection quality and return potential on the basis of the $N = 5,000$ yearly returns. By using downside risk measures, we are able to appropriately quantify the protection quality of any given strategy. In particular, we

[9] Benartzi and Thaler (1995) provide evidence that many institutional and private investors use a one-year investment horizon. Therefore, we focus on this time period in our simulation analysis.

[10] As a robustness check, we also implement a circular block-bootstrap in order to avoid an underweighting at the beginning and the end of the original time series. Our results remain qualitatively the same.

[11] We also run the simulations using the original volatility or the Leland (1985) modification in equation (7a). The simulation results are not sensitive to these changes.

compute lower partial moments (LPMs), which take only negative deviations from the pre-specified target return, denoted as τ, into account. As our analysis focuses on preserving the invested capital, the target return is set equal to zero ($\tau = 0$), and we penalize negative deviations from this target with an exponent of 0 (LPM0) and 1 (LPM1). LPM0 and LPM1 represent the shortfall probability and the expected shortfall, respectively (Harlow and Rao (1989), Eling and Schuhmacher (2007)).[12] We also report the empirical VaR (using a confidence level of $a = 95\%$), the skewness, and the minimum annual return as alternative downside risk measures.

Following Bertrand and Prigent (2011), we use the Omega measure (Shadwick and Keating (2002)) for our systematic comparison of the different portfolio insurance strategies along the distinguishing characteristics as discussed in Section 5.2.6 (comparisons 1-4). The Omega ratio represents a special case of the more general Kappa measure (Kaplan and Knowles (2004)). It is characterized by the ratio of gains to losses relative to a predefined target return required by the investor:

$$\Omega(\tau) = \frac{\int_{\tau}^{\infty}(1 - F(r))dr}{\int_{-\infty}^{\tau} F(r)dr}, \tag{14}$$

where $F(.)$ denotes the cumulative distribution function of returns r, and τ represents the target return (or threshold). The Omega ratio takes account of the entire return distribution, does not require any parametric assumptions, and it can easily be computed by dividing the higher partial moment of degree 1 (HPM1) by the lower partial moment of the same degree (LPM1). In order to get a reasonable ratio of gains to losses, the target return should be lower than the expected portfolio return. In the context of portfolio insurance strategies, the threshold must also be higher than the protection level. Following Bertrand and Prigent (2011), we set the threshold for the Omega ratio to 1%, 2%, 3%, and 4% per year, respectively.

5.3.3 Testing for Statistical Significance

The starting point for our statistical significance tests are the $N = 5,000$ annual portfolio returns for each simulated protection strategy (based on the block-bootstrap approach from the first step). For all comparisons 1-4 formulated in Section 5.2.6, we evaluate the difference between the Omega ratios of a pair of strategies A and B:

$$\Delta_{\Omega} = \Omega_A(\tau) - \Omega_B(\tau). \tag{15}$$

[12] In the specific case of a zero target return, the LPM0 measure indicates the loss probability and the LPM1 measure the expected loss.

The objective is to test whether there is a significant difference between the Omega ratios of the two compared strategies. Formally, we test:

$$H_0 : \Delta_\Omega = 0 \qquad \text{against} \qquad H_1 : \Delta_\Omega \neq 0. \tag{16}$$

Given the $N = 5,000$ annual returns for each strategy, an appropriate point estimator for the difference in equation (15) can be calculated:

$$\hat{\Delta}_\Omega = \hat{\Omega}_A(\tau) - \hat{\Omega}_B(\tau). \tag{17}$$

However, a statistical test requires the distribution of $\hat{\Delta}_\Omega$. Therefore, we implement a further bootstrap using the $N = 5,000$ annual returns from the block-bootstrap in the first step. Within this second step bootstrap, we again create 5,000 bootstrap resamples ($N_B = 5,000$), each consisting of 5,000 annual returns. For each of the $N_B = 5,000$ bootstrap resamples, we compute in the same way as for the point estimator in equation (17). Furthermore, we define:

$$\Delta^*_{\Omega[1]} \leq \Delta^*_{\Omega[2]} \leq \ldots \leq \Delta^*_{\Omega[N_B]} \tag{18}$$

as the ordered series of the Omega differences. Based on this series a confidence interval:

$$CI = [\xi^*_{low}; \xi^*_{high}] \tag{19}$$

can be constructed, where:

$$\xi^*_{low} = \Delta^*_{\Omega[\frac{\alpha}{2}\cdot N_B]} \qquad \text{and} \tag{20a}$$

$$\xi^*_{high} = \Delta^*_{\Omega[(1-\frac{\alpha}{2})\cdot N_B]}. \tag{20b}$$

The null hypothesis H_0 is rejected at significance level α if $0 \notin CI$. The key benefit of applying a block-bootstrap procedure for a comparison of portfolio insurance strategies is that all time-series properties are taken into account. A caveat, however, is that a selection of blocks of daily returns with replacement leads to overlapping effects, and thus the $N = 5,000$ generated annual returns from the first step are not fully independent. Therefore, the standard bootstrap as proposed by Efron (1979) is not appropriate.[13] We therefore use the stationary bootstrap approach suggested by Politis

[13] A potential solution could be the block-bootstrap approach, as proposed by Künsch (1989) and Liu and Singh (1992). However, a problem is that the resampled time series need not necessarily be stationary

and Romano (1994), where the block length in the bootstrap is stochastic and obeys a geometric distribution. Figure I graphically summarizes our methodology, which consists of a traditional block-bootstrap for the simulation of the annual returns of the respective portfolio insurance strategy in the first step (Panel A) and the stationary bootstrap of Politis and Romano (1994) for conducting the hypothesis tests in the second step (Panel B).

Specifically, let $R_i = (R_{A_i}, R_{B_i})$ be a pair of the annual returns of strategy A and B in the second step of our analysis. The starting point are the N pairs of annual returns, denoted as $(R_1, \ldots, R_N)$, where we calculate the point estimator $\hat{\Delta}_\Omega = (R_1, \ldots, R_N)$. Similarly, we calculate the $\hat{\Delta}^*_\Omega$ values as $\hat{\Delta}^*_\Omega = \Delta_\Omega(R^*_1, \ldots, R^*_N)$, where the R^*_i values, with $i \in \{1, \ldots, N\}$, are results from the stationary bootstrap. Instead of resampling single elements or blocks of fixed length with replacement from the series $(R_1, \ldots, R_N)$, the stationary bootstrap selects blocks of data with a variable block length: $(R^*_1, \ldots, R^*_N) = B_{I_1,L_1}, B_{I_2,L_2}, \ldots$ The first data block, denoted as B_{I_1,L_1}, starts at index I_1 and has L_1 elements. In general, we have:

$$B_{I_i,L_i} = \{R_{I_i}, R_{I_i+1}, \ldots, R_{I_i+L_i-1}\}. \tag{21}$$

The starting indices $I_1, I_2, \ldots$ are random variables drawn from a discrete uniform distribution. The block lengths L_i are random variables described by a geometric distribution:

$$P(L_i = m) = (1-p)^{m-1} \cdot p \qquad \text{for } m \in \mathbb{N}. \tag{22}$$

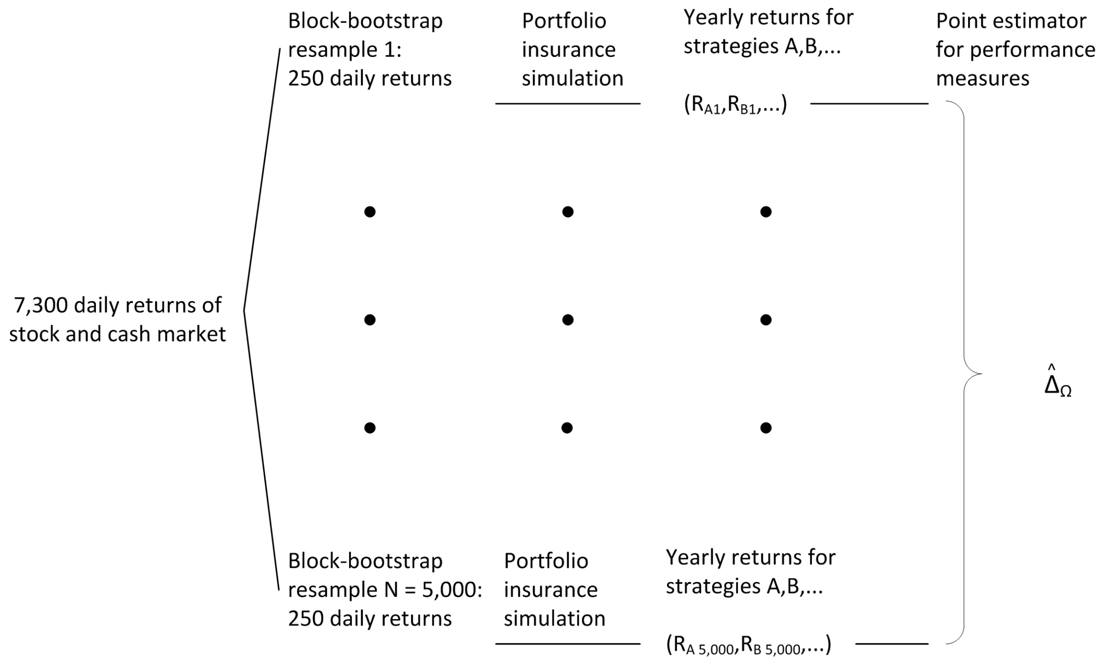
When building the blocks as described, the index of the next element (j) that has to be selected could be greater than $N (j > N)$. In this case, R_j is defined to be R_i, where $i = (j \bmod N)$ and $R_0 = R_N$. Politis and Romano (1994) emphasize that this 'wrapping the data around in a circle' (such that R_1 follows R_N) is important to achieve stationarity of the resampled time series. The geometric distribution, as shown in equation (22), has an expected value of $(1/p)$ which represents the mean length of our data blocks. Therefore, specifying the probability p allows controlling the mean length of the selected data blocks at the same time. In our implementation of the stationary bootstrap, we use the algorithm developed by Politis and White (2004) to determine the 'optimal' mean block length. This algorithm implicitly solves for the optimal p. Patton et al. (2009) suggest minor error corrections to this algorithm, and we further incorporate their proposed changes.

The stationary bootstrap test can be implemented in a fairly straightforward manner. In a first step, generate an integer random number i in the interval $\{1, \ldots, N\}$ and select

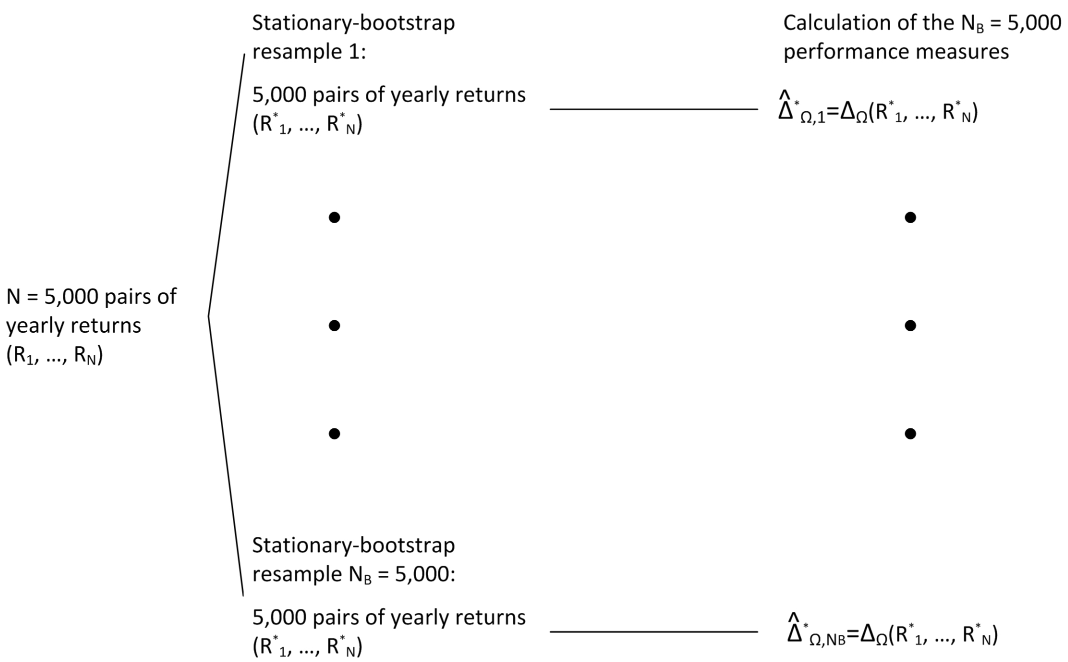
---

although the original series of elements fulfill this assumption. Politis and Romano's (1994) stationary bootstrap technique alleviates this problem.

Figure I – The Historical Bootstrap Set-Up



(A) Block-Bootstrap Simulation



(B) Stationary Bootstrap Simulation

the element R_i. This number represents the first element in the first block: $R_1^* = R_{I_1} = R_i$. Then generate another real value random number in the interval $[0;1[$. If this generated random number is greater than the specified probability p, select the next element R_{i+1}, set $R_2^* = R_{I_1+1} = R_{i+1}$, and generate a new real value random number in the interval $[0;1[$. If the selected index j is greater than N, start again with the first element R_1, and so on ('wrapping the data around a circle') until the first selected block has been built. However, if the generated real value random number is smaller than p, building up the first block B_{I_1,L_1} ends. We then generate a new integer random number i in the interval $\{1,\ldots,N\}$, select an element R_i, and build the next block B_{I_2,L_2}. The process stops once we have N resampled pairs of returns $(R_1^*,\ldots,R_N^*)$, where we can calculate $\hat{\Delta}_\Omega^* = \Delta_\Omega(R_1^*,\ldots,R_N^*)$. This procedure is repeated N_B times until we have N_B values of $\hat{\Delta}_\Omega^*$ representing the distribution of $\hat{\Delta}_\Omega$.

5.4 Historical Simulation Results

5.4.1 Main Simulation Results

The results of our bootstrap simulations are presented in Panel A of Table II. As expected, all portfolio insurance strategies generate positive skewness, which is in contrast to the negative skewness of the aggregate stock market return series. On average, the stop-loss strategy provides the highest mean annual return, but also the highest risk. In terms of average annual returns, both the synthetic put and the dynamic VaR-strategy also dominate the CPPI and the TIPP strategy. In contrast, the CPPI strategy and the TIPP strategy dominate all other strategies in terms of their risks, as indicated by the LPM0 and the LPM1 measure, the empirical VaR (at the 95% confidence level), and the minimum annual return (as measures for protection quality). Therefore, the resulting question of interest is which portfolio insurance strategy leads to the highest risk-adjusted performance.

In order to implement our systematic strategy comparison, we evaluate the different portfolio insurance strategies based on their Omega ratio by reporting statistical significance levels. The Omega ratios are shown in Panel A of Table II for threshold returns ranging from 1% to 4%. Most important, a bootstrap-based hypothesis test as described in Section 5.3.3 is applied to statistically test the differences in the Omega ratios for each pair of strategies (see Section 5.2.6). An average block length of 3 is chosen to conduct all hypothesis tests.[14] Confidence intervals are evaluated at the 1%, 5%, and 10% level of statistical significance, respectively. Panel B of Table II reports

[14] In order to determine the optimal block length of our bootstrap-based hypothesis tests, we apply Politis and White's (2004) algorithm and also incorporate the error corrections suggested by Patton et al. (2009). All calculated values vary around 2, and we choose an average block length of 3 in all cases in order to get conservative results. A block length of 2 delivers almost the same results.

the confidence interval with the highest level of significance. If the confidence interval based on a 10% significance level still contains zero, there are no significant differences in the Omega ratios for the respective pair of strategies. In contrast, if both reported interval boundaries are positive (negative), one strategy features a significantly higher (lower) Omega ratio than the other at the reported level of statistical significance.

Comparing the simple stop-loss trading rule with the CPPI strategy (comparison 1), we observe a clear dominance of the CPPI strategy. However, the difference in the Omega ratio in Panel B of Table II shrinks with increasing threshold returns, which can be attributed to the lower return potential of the CPPI strategy in comparison to the stop-loss strategy. With a threshold return of $\tau = 4\%$, the difference in the Omega ratios is no longer statistically significant. This observation based on historical data contrasts the findings in Benninga (1990), who reports a dominance of the stop-loss strategy in terms of risk-adjusted returns. However, he does not use downside risk measures and conducts Monte Carlo simulations, potentially neglecting the properties of financial markets data (e.g., heteroscedasticity and fat tails).

In a second step, we compare the CPPI strategy and the TIPP strategy (comparison 2). Panel B of Table II reveals a statistically significant dominance of the TIPP strategy in terms of the Omega ratio with threshold returns up to 2%. Using a threshold return of 3%, the difference is no longer statistically significant. However, using a threshold return of 4%, the CPPI strategy provides a significantly higher Omega ratio compared to the TIPP strategy. This reversal is explained by the tradeoff between downside protection and return potential. Although the TIPP strategy exhibits a slightly superior downside protection (as indicated by the LPM0 and the LPM1 measure), it also provides a lower return potential. In order to assess whether these properties are specific for the chosen multiplier of $m = 5$ in the base-case scenario, we run additional bootstrap simulations and compare the CPPI strategy and the TIPP strategy with multipliers of $m = 3$, $m = 7$, and $m = 9$. The results are summarized in Table III; they generally confirm our earlier findings in Table II. While the TIPP strategy exhibits slightly superior risk measures, the CPPI strategy dominates in terms of annual mean returns. With high multipliers of $m = 7$ and $m = 9$, the mean return of the TIPP strategy does not increase notably, and the CPPI strategy generates higher Omega ratios than the TIPP strategy when using threshold returns of 3% and 4%. The limited return potential of the TIPP strategy is a result of its property to protect the cumulated wealth. Given the continuous ratcheting up of the floor to the highest portfolio value, this increased floor is more often violated, implying a full cash investment until the end of the investment horizon and lower portfolio returns. It is expected that this effect will reinforce itself with increasing investment horizons. Overall, these results reconfirm Choie and Seff's (1989) general criticism of the TIPP strategy.

Table II – Historical Simulation Results

This table shows the results of historical simulations using the base-case specifications. The data consists of daily returns for the German stock market index DAX and money market rates from the German Bundesbank (average values of the 1-month Frankfurt interbank rate; middle rate) over the sample period from January 1981 to December 2011. Volatility and risk premium estimates (if necessary) are taken from the year prior to the implementation of the strategy. In order to capture the time series characteristics of the return series, we implement a common block-bootstrap approach drawing 250 subsequently following pairs of daily stock and money market returns with replacement. In particular, selecting $N = 5,000$ coherent data blocks with replacement and applying the different portfolio insurance strategies in each draw provides us with 5,000 yearly returns for each strategy. Panel A presents the return, risk, and performance characteristics of the different portfolio insurance strategies. $m = 5$ denotes the multiplier of the CPPI and the TIPP strategy, and $\alpha = 0.90$ is the confidence level of the dynamic VaR-strategy. In addition to the mean annual returns, the strategies' protection quality is captured by using downside risk measures. Lower partial moments (LPMs) only take negative deviations from the target return (τ) into account. The target return is set equal to zero ($\tau = 0$), and we penalize negative deviations from this target with an exponent of 0 (LPM0) and 1 (LPM1). LPM0 and LPM1 represent the shortfall probability and the expected shortfall, respectively. In addition, the table reports the empirical VaR (using a confidence level of $\alpha = 95\%$), the skewness, and the minimum annual return as alternative downside risk measures. The Omega measure, defined as the ratio of gains to losses relative to a target return, is used for our systematic comparison of the different portfolio insurance strategies along the distinguishing characteristics as discussed in Section 5.2.6 (comparisons 1-4 in Panel B). The threshold returns for the Omega ratio are set to 1%, 2%, 3%, and 4% per year, respectively. Panel B reports the confidence intervals for significance tests of the difference between the Omega ratios of each pair of strategies. In order to derive the confidence intervals, we implement a stationary bootstrap of Politis and Romano (1994) using the $N = 5,000$ annual returns from the common block-bootstrap in the first step. Within this second step bootstrap, we again create $N_B = 5,000$ bootstrap resamples, each consisting of 5,000 annual returns (see Figure I). The stationary bootstrap approach suggested by Politis and Romano (1994) is used for all hypothesis tests in the second step, where the optimal mean block length is determined as suggested by Patton et al. (2009). The confidence intervals are evaluated at the 1%, 5%, and 10% level of significance, and Panel B reports the confidence interval with the highest level of significance. If the confidence interval based on a 10% significance level still contains zero, there are no significant differences in the Omega ratios of the respective pair of strategies. In contrast, if both reported interval boundaries are positive (negative), one strategy features a significantly higher (lower) Omega ratio than the other at the reported level of statistical significance. *, **, and *** denotes statistical significance at the 10%, 5%, and 1% level, respectively.

Panel A: Return, Risk, and Performance of Portfolio Insurance Strategies

	Stop-Loss	Synthetic Put	CPPI ($m = 5$)	TIPP ($m = 5$)	Dynamic VaR ($\alpha = 0.90$)	Stock Market
Mean return p.a.	8.97%	8.19%	5.82%	4.88%	7.31%	11.36%
LPM0 ($\tau = 0\%$)	68.10%	28.94%	6.98%	3.48%	28.54%	31.84%
LPM1 ($\tau = 0\%$)	1.06%	0.55%	0.04%	0.02%	0.40%	5.41%
VaR (95%)	−3.73%	−3.02%	−0.12%	0.24%	−2.36%	−30.28%
Skewness	1.58	1.78	2.07	0.68	2.32	−0.08
Minimum return	−9.80%	−8.66%	−2.21%	−2.21%	−9.15%	−57.32%
$\Omega(\tau = 1\%)$	5.57	9.14	25.03	39.85	9.32	2.81
$\Omega(\tau = 2\%)$	3.87	5.73	8.89	11.27	5.42	2.54
$\Omega(\tau = 3\%)$	2.92	3.91	4.24	4.33	3.51	2.31
$\Omega(\tau = 4\%)$	2.31	2.83	2.35	1.91	2.46	2.09

continued

Table II – *Continued*

Panel B: Statistical Significance					
	Comparison 1	Comparison 2	Comparison 3	Comparison 4	
	CPPI versus Stop-Loss	CPPI versus TIPP	CPPI versus Synthetic Put	Dynamic VaR versus Synthetic Put	Dynamic VaR versus CPPI
$\Omega(\tau = 1\%)$	[17.12;22.18] ***	[−18.91;−11.39] ***	[13.75;18.38] ***	[−0.37;0.73]	[−18.36;−13.41] ***
$\Omega(\tau = 2\%)$	[4.37;5.71] ***	[−3.08;−1.73] ***	[2.64;3.75] ***	[−0.57;−0.05] *	[−4.15;−2.87] ***
$\Omega(\tau = 3\%)$	[1.06;1.60] ***	[−0.25;0.06]	[0.14;0.53] ***	[−0.66;−0.14] ***	[−1.00;−0.47] ***
$\Omega(\tau = 3\%)$	[−0.05;0.14]	[0.32;0.57] ***	[−0.58;0.37] ***	[−0.54;−0.21] ***	[0.01;0.19] *

Table III – Hist. Simulation Results for CPPI and TIPP Strategies with Different Multipliers
With respect to the measures of interest, this table shows the return, risk, and performance values of historical simulations for the CPPI strategy and TIPP strategy with different choices for the multiplier m (see equation 9). The data consists of daily returns for the German stock market index DAX and money market rates from the German Bundesbank (average values of the 1-month Frankfurt interbank rate; middle rate) over the sample period from January 1981 to December 2011. In order to capture the time series characteristics of the return series, we implement a common block-bootstrap approach by drawing 250 subsequently following pairs of daily stock and money market returns with replacement. In particular, selecting $N = 5,000$ coherent data blocks with replacement and applying the different portfolio insurance strategies in each draw provides us with 5,000 yearly returns for each strategy. In addition to the mean annual returns, the strategies' protection quality is captured by using downside risk measures. Lower partial moments (LPMs) only take negative deviations from the target return (τ) into account. The target return is set equal to zero ($\tau = 0$), and we penalize negative deviations from this target with an exponent of 0 (LPM0) and 1 (LPM1). LPM0 and LPM1 represent the shortfall probability and the expected shortfall, respectively. In addition, the table reports the values of the empirical VaR (using a confidence level of $\alpha = 95\%$), the skewness, and the minimum annual return as alternative downside risk measures. The Omega measure is defined as the ratio of gains to losses relative to a target return, where the threshold returns are set to 1%, 2%, 3%, and 4% per year, respectively.

	CPPI ($m = 3$)	TIPP ($m = 3$)	CPPI ($m = 7$)	TIPP ($m = 7$)	CPPI ($m = 9$)	TIPP ($m = 9$)
Mean return p.a.	5.30%	4.84%	6.61%	4.97%	7.10%	5.05%
LPM0 ($\tau = 0\%$)	1.98%	1.40%	15.02%	6.54%	28.06%	10.74%
LPM1 ($\tau = 0\%$)	0.01%	0.01%	0.07%	0.04%	0.17%	0.12%
VaR (95%)	0.42%	0.57%	0.44%	−0.17%	−1.34%	−0.68%
Skewness	0.96	0.34	2.08	0.92	2.18	1.03
Minimum return	−1.35%	−1.34%	−2.76%	−2.76%	−3.85%	−3.85%
$\Omega(\tau = 1\%)$	50.36	70.55	17.57	25.36	11.95	15.09
$\Omega(\tau = 2\%)$	12.43	15.21	7.19	8.01	5.89	6.12
$\Omega(\tau = 3\%)$	4.96	5.30	3.94	3.52	3.59	3.09
$\Omega(\tau = 4\%)$	2.35	2.06	2.48	1.79	2.44	1.73

In a third step, we compare the synthetic put strategy and the CPPI strategy (comparison 3). As shown in Panel A of Table II, the synthetic put strategy dominates the CPPI strategy in terms of mean annual returns. However, the synthetic put strategy provides worse downside protection, as indicated by the values of the LPM0 and LPM1 measures, the empirical VaR, and the minimum return. Based on the performance of the Omega ratio, we observe a statistically significant dominance of the CPPI strategy against the synthetic put strategy for threshold returns up to 3% in Panel B of Table II. In contrast, using a threshold return of 4%, the synthetic put strategy is superior to the CPPI strategy in terms of Omega ratios at the 1% significance level. One explanation for this reversal is the choice of the multiplier. As shown in Table III, the return potential of the CPPI strategy increases with higher multipliers m. Moreover, the relatively poor protection quality of the synthetic put strategy may be explained by estimation risk for stock market volatility. Both issues are addressed in Section 5.4.2.

In a fourth step, we evaluate the dynamic VaR-strategy against both the synthetic put and the CPPI strategy (comparison 4). The dynamic-VaR strategy shows a lower return potential than the synthetic put strategy, but it also generates slightly superior values for the downside risk measures in Panel A of Table II. Moreover, based on the Omega ratio, we find a statistically significant dominance of the synthetic put strategy for threshold returns from 2% to 4%. The results are opposite for comparisons of the VaR-strategy with the CPPI strategy. While the VaR-strategy provides higher mean annual returns, the CPPI strategy features superior downside protection properties. For threshold returns up to 3%, the Omega ratio of the CPPI strategy is significantly higher than that of the dynamic VaR-strategy.

Table IV – Hist. Simulation Results for the Dynamic VaR-Strategy with Diff. Confidence Levels
With respect to the measures of interest, this table shows the return, risk, and performance values of historical simulations for the dynamic VaR-strategy with different choices for the confidence level α (equation 13). The data consists of daily returns for the German stock market index DAX and money market rates from the German Bundesbank (average values of the 1-month Frankfurt interbank rate; middle rate) over the sample period from January 1981 to December 2011. Volatility and risk premium estimates are taken from the year prior to the implementation of the strategy. In order to capture the time series characteristics of the return series, we implement a common block-bootstrap approach by drawing 250 subsequently following pairs of daily stock and money market returns with replacement. In particular, selecting $N = 5,000$ coherent data blocks with replacement and applying the different portfolio insurance strategies in each draw provides us with 5,000 yearly returns for each strategy. In addition to the mean annual returns, the strategies' protection quality is captured by using downside risk measures. Lower partial moments (LPMs) only take negative deviations from the target return (τ) into account. The target return is set equal to zero ($\tau = 0$), and we penalize negative deviations from this target with an exponent of 0 (LPM0) and 1 (LPM1). LPM0 and LPM1 represent the shortfall probability and the expected shortfall, respectively. In addition, the table reports the values of the empirical VaR (using a confidence level of $\alpha = 95\%$), the skewness, and the minimum annual return as alternative downside risk measures. The Omega measure is defined as the ratio of gains to losses relative to a target return, where the threshold returns are set to 1%, 2%, 3%, and 4% per year, respectively.

	Confidence Level α			
	$\alpha = 0.85$	$\alpha = 0.90$	$\alpha = 0.95$	$\alpha = 0.99$
Mean return p.a.	7.21%	6.97%	6.80%	6.13%
LPM0 ($\tau = 0\%$)	32.22%	28.60%	21.44%	10.44%
LPM1 ($\tau = 0\%$)	0.38%	0.39%	0.28%	0.15%
VaR (95%)	−2.31%	−2.37%	−1.80%	−0.55%
Skewness	2.37	2.33	2.33	2.24
Minimum return	−7.95%	−9.15%	−8.70%	−9.82%
$\Omega(\tau = 1\%)$	5.10	8.96	11.00	15.39
$\Omega(\tau = 2\%)$	2.27	5.16	5.90	6.93
$\Omega(\tau = 3\%)$	3.43	3.32	3.61	3.78
$\Omega(\tau = 4\%)$	2.40	2.30	2.40	2.31

A caveat is that the dynamic VaR-strategy is implemented by using a fixed confidence level of $\alpha = 90\%$ (see equation 11), and it is not obvious whether this value is the optimal choice. In order to evaluate the influence of the confidence level, we repeat our bootstrap simulations for varying values of the parameter α in a dynamic VaR-strategy, ranging from 0.85 to 0.99. The results are summarized in Table IV. As expected, a reduction in the level of significance from 0.15 to 0.01 leads to superior downside protection, but also to a lower return potential. While a dynamic VaR-strategy with $\alpha = 0.85$ boasts the highest mean annual return, the corresponding downside protection tends to be worse compared to the specifications with higher significance levels. The dynamic VaR-strategy based on $\alpha = 0.95$ appears particularly interesting. Comparing this strategy with the $\alpha = 0.99$ specification, a higher mean annual return potential is observable. At the same time, this strategy dominates the $\alpha = 0.85$ and $\alpha = 0.90$ specifications in terms of the LPM0 and the LPM1 measure as well as the empirical VaR. In addition, based on the Omega ratios, a dominance of the $\alpha = 0.95$ specification against $\alpha = 0.90$ is observable for all threshold returns ranging from 1% to 4%.

5.4.2 Robustness Tests

In this section, we conduct a series of robustness tests. Given the results in Tables III and IV, we start by repeating our bootstrap simulations with alternative parameterizations of the CPPI strategy and the dynamic VaR-strategy. The confidence level for the dynamic VaR-strategy is set to $\alpha = 0.95$, and the CPPI strategy is implemented with a multiplier of $m = 9$.[15] The results are presented in Table V. With this more aggressive multiplier (compared to $m = 5$), the CPPI strategy again strongly dominates the simple stop-loss trading rule in terms of Omega ratios (comparison 1). The difference is statistically significant for all threshold returns ranging from 1% to 4% (Panel B of Table V). The comparison of the CPPI and the TIPP strategy using a multiplier of $m = 9$ also reconfirms our earlier results (comparison 2). Again, the CPPI strategy generates significantly higher Omega ratios compared to the TIPP strategy using threshold returns of 3% and 4%.

Comparing the CPPI strategy with the synthetic put strategy (comparison 3), our earlier results change. With the higher multiplier of $m = 9$, the CPPI strategy is dominated by the synthetic put strategy in terms of Omega ratios even by using a threshold return of 3%. While the synthetic put strategy provides higher mean returns than the CPPI strategy, the latter still offers superior downside protection.[16] A comparison of the dynamic VaR-strategy with the synthetic put strategy again confirms earlier findings (comparison 4). The dynamic VaR-strategy exhibits superior values for the LPM0 and the LPM1 measure as well as the empirical VaR, but it also shows a lower return potential compared to the synthetic put strategy. In terms of the Omega ratio, the VaR-strategy significantly dominates the synthetic put strategy only for low thresholds

---

[15] In order to enable a comparison between the CPPI and the TIPP strategy, we also set the multiplier for the TIPP strategy to $m = 9$. The deviations in the entries for both the CPPI and the TIPP strategy based on $m = 9$ in Tables III and V are attributable to the path dependency of each portfolio insurance strategy, respectively. Similar deviations are observable for the stop-loss strategy and the dynamic VaR-strategy across tables. However, all deviations are very small, indicating that our bootstrap simulations exhibit sufficiently accurate convergence properties.

[16] Our results differ slightly from the findings in Bertrand and Prigent (2011), where the CPPI strategy dominates the synthetic put strategy under a full capital protection in terms of the Omega measure for all threshold returns, ranging from 1% to 4%.

Table V – Robustness Checks for Historical Simulation Results

This table shows the results of historical simulations using alternative strategy specifications. In particular, in contrast to the base-case specification in Table II, a multiplier of $m = 9$ (instead of $m = 5$) is used for both the CPPI and the TIPP strategy and a confidence level of $\alpha = 0.95$ (instead of $\alpha = 0.90$) for the dynamic VaR-strategy. The data consists of daily returns for the German stock market index DAX and money market rates from the German Bundesbank (average values of the 1-month Frankfurt interbank rate; middle rate) over the sample period from January 1981 to December 2011. Volatility and risk premium estimates (if necessary) are taken from the year prior to the implementation of the strategy. In order to capture the time series characteristics of the return series, we implement a common block-bootstrap approach by drawing 250 subsequently following pairs of daily stock and money market returns with replacement. In particular, selecting $N = 5,000$ coherent data blocks with replacement and applying the different portfolio insurance strategies in each draw provides us with 5,000 yearly returns for each strategy. Panel A presents the return, risk, and performance characteristics of the different portfolio insurance strategies. In addition to the mean annual returns, the strategies' protection quality is captured by using downside risk measures. Lower partial moments (LPMs) only take negative deviations from the target return (τ) into account. The target return is set equal to zero ($\tau = 0$), and we penalize negative deviations from this target with an exponent of 0 (LPM0) and 1 (LPM1). LPM0 and LPM1 represent the shortfall probability and the expected shortfall, respectively. In addition, the table reports the values of the empirical VaR (using a confidence level of $\alpha = 95\%$), the skewness, and the minimum annual return as alternative downside risk measures. The Omega measure, defined as the ratio of gains to losses relative to a target return, is used for our systematic comparison of the different portfolio insurance strategies along the distinguishing characteristics as discussed in section 5.2.6 (comparisons 1-4 in Panel B). The threshold returns for the Omega ratio are set to 1%, 2%, 3%, and 4% per year, respectively. Panel B reports the confidence intervals for significance tests of the difference between the Omega ratios of each pair of strategies. In order to derive the confidence intervals, we implement a stationary bootstrap of Politis and Romano (1994) using the $N = 5,000$ annual returns from the common block-bootstrap in the first step. Within this second step bootstrap, we again create $N_B = 5,000$ bootstrap resamples, each consisting of 5,000 annual returns (see Figure I). The stationary bootstrap approach suggested by Politis and Romano (1994) is used for all hypothesis tests in the second step, where the optimal mean block length is determined as suggested by Patton et al. (2009). The confidence intervals are evaluated at the 1%, 5%, and 10% level of significance, and Panel B reports the confidence interval with the highest level of significance. If the confidence interval based on a 10% significance level still contains zero, there are no significant differences in the Omega ratios of the respective pair of strategies. In contrast, if both reported interval boundaries are positive (negative), one strategy features a significantly higher (lower) Omega ratio than the other at the reported level of statistical significance. ***, **, and * denotes statistical significance at the 1%, 5%, and 10% level, respectively.

Panel A: Return, Risk, and Performance of Portfolio Insurance Strategies

	Stop-Loss	Synthetic Put	CPPI ($m = 9$)	TIPP ($m = 9$)	Dynamic VaR ($\alpha = 0.95$)	Stock Market
Mean return p.a.	8.78%	8.01%	6.96%	5.03%	6.91%	11.09%
LPM0 ($\tau = 0\%$)	68.40%	28.26%	24.42%	10.66%	21.46%	31.50%
LPM1 ($\tau = 0\%$)	1.08%	0.53%	0.16%	0.10%	0.27%	5.35%
VaR (95%)	−3.97%	−2.96%	−1.21%	−0.67%	−1.76%	−29.40%
Skewness	1.56	1.76	2.20	1.04	2.36	−0.10
Minimum return	−10.62%	−8.63%	−3.85%	−3.77%	−9.12%	−58.81%
$\Omega(\tau = 1\%)$	5.42	9.17	12.15	15.85	11.26	2.78
$\Omega(\tau = 2\%)$	3.77	5.70	5.90	6.30	6.04	2.51
$\Omega(\tau = 3\%)$	2.84	3.86	3.56	3.12	3.70	2.28
$\Omega(\tau = 4\%)$	2.25	2.77	2.40	1.72	2.47	2.06

continued

Table V – *Continued*

Panel B: Statistical Significance					
	Comparison 1	Comparison 2	Comparison 3	Comparison 4	
	CPPI versus Stop-Loss	CPPI versus TIPP	CPPI versus Synthetic Put	Dynamic VaR versus Synthetic Put	Dynamic VaR versus CPPI
$\Omega(\tau = 1\%)$	[5.90;7.62] ***	[−4.97;−2.63] ***	[2.33;3.71] ***	[1.13;3.13] ***	[−1.63;−0.13] *
$\Omega(\tau = 2\%)$	[1.82;2.48] ***	[−0.82;−0.02] ***	[0.00;0.41] **	[0.01;0.68] **	[−0.16;0.44]
$\Omega(\tau = 3\%)$	[0.54;0.91] ***	[0.23;0.65] ***	[−0.44;−0.15] ***	[−0.32;0.00]	[−0.03;0.30]
$\Omega(\tau = 3\%)$	[0.02;0.27] ***	[0.53;0.82] ***	[−0.49;−0.28] ***	[−0.47;−0.15] ***	[−0.03;0.18]

of 1% and 2%. Comparing the Omega ratios of the dynamic VaR-strategy and the CPPI strategy, slightly better results are observable for the dynamic VaR-strategy. In fact, the advantage of the CPPI strategy against the dynamic VaR-strategy seems to be attributable to its lower risk, and this advantage is more pronounced when a lower multiplier for the CPPI strategy is chosen.

Table VI – Hist. Simulation Results for the Dynamic VaR-Strategy with Diff. Confidence Levels
With respect to the measures of interest, this table shows return, risk, and performance values of historical simulations for the dynamic VaR-strategy with different choices for the confidence level α (equation 13) and the synthetic put strategy assuming that estimation risk is ignored. In contrast to all previous simulations, where the volatility and risk premium estimates are taken from the year prior to the implementation of the strategy, the realized volatility and the realized risk premium in a given year are used. The data consists of daily returns for the German stock market index DAX and money market rates from the German Bundesbank (average values of the 1-month Frankfurt interbank rate; middle rate) over the sample period from January 1981 to December 2011. In order to capture the time series characteristics of the return series, we implement a common block-bootstrap approach by drawing 250 subsequently following pairs of daily stock and money market returns with replacement. In particular, selecting $N = 5,000$ coherent data blocks with replacement and applying the different portfolio insurance strategies in each draw provides us with 5,000 yearly returns for each strategy. In addition to the mean annual returns, the strategies' protection quality is captured by using downside risk measures. Lower partial moments (LPMs) only take negative deviations from the target return (τ) into account. The target return is set equal to zero ($\tau = 0$), and we penalize negative deviations from this target with an exponent of 0 (LPM0) and 1 (LPM1). LPM0 and LPM1 represent the shortfall probability and the expected shortfall, respectively. Moreover, the table reports the values of the empirical VaR (using a confidence level of $\alpha = 95\%$), the skewness, and the minimum annual return as alternative downside risk measures. The Omega measure is defined as the ratio of gains to losses relative to a target return, where the threshold returns are set to 1%, 2%, 3%, and 4% per year, respectively.

	Confidence Level α				Synthetic
	$\alpha = 0.85$	$\alpha = 0.90$	$\alpha = 0.95$	$\alpha = 0.99$	Put
Mean return p.a.	4.96%	5.79%	6.86%	7.69	7.89%
LPM0 ($\tau = 0\%$)	27.12%	17.42%	8.60%	2.20	21.90%
LPM1 ($\tau = 0\%$)	0.27%	0.21%	0.10%	0.01	0.20%
VaR (95%)	−1.70%	−1.49%	−0.53%	0.37	−1.32%
Skewness	2.40	2.18	1.97	2.53	1.68
Minimum return	−7.56%	−7.17%	−6.67%	−4.50	−4.12%
$\Omega(\tau = 1\%)$	6.99	10.63	22.43	84.19	14.35
$\Omega(\tau = 2\%)$	3.49	5.07	9.47	22.89	7.13
$\Omega(\tau = 3\%)$	2.09	2.88	4.82	9.24	4.35
$\Omega(\tau = 4\%)$	1.39	1.85	2.85	4.70	2.96

In contrast to all other strategies, the synthetic put strategy and the dynamic VaR-strategy are both dependent on the parameter forecasts. Comparing the LPM0 measures with a given parameter α for all different specifications of the dynamic VaR-strategy in Tables II, IV, and V, respectively, substantial differences emerge. In all cases, the LPM0 measure of the dynamic VaR-strategy indicates that the protection level is more often violated than expected given the confidence level α. Furthermore, the synthetic put strategy exhibits a worse capital protection compared to other forecast-free strategies, such as the CPPI strategy and the TIPP strategy. While deviations from the assumption of normally distributed stock market returns may be one plausible explanation, the limited prediction quality of our volatility and risk premium estimators may offer another explanation. In order to shed light on this latter issue, we repeat our simulations for the dynamic VaR-strategy and the synthetic put strategy. In contrast to our previous simulations, where the volatility and risk premium estimates are taken from the year prior to the implementation of the strategy, we use the realized volatility and the realized risk premium in any given year. The results are shown in Table VI. Comparing the LPM0 measures of the dynamic VaR-strategy with and without estimation risk in Table IV and Table VI, respectively, we observe a substantial reduction in the LPM0 value towards the expected value as indicated by the confidence level α. For example, based on $\alpha = 0.99$, the value of the LPM0 measure decreases from 10.44% to 2.20% when estimation risk is fully eliminated. However, the expected value of 1% still cannot be reached. This effect is observable for all levels of significance.

Another important effect is revealed when the dynamic VaR-strategy is compared for increasing confidence levels α (ranging from $\alpha = 0.85$ to $\alpha = 0.99$). On the one hand, protection quality is improved when the confidence level is increased. This expected result is indicated by the values of the LPM0 and LPM1 measures, the empirical VaR, and the minimum return. On the other hand, we also observe an increase in the mean annual return with increasing confidence levels (from 4.96% for $\alpha = 0.85$ to 7.69% for $\alpha = 0.99$). A more conservative dynamic VaR-strategy not only reduces risk, but it also enhances the return potential under 'perfect' forecasts.[17] These results suggest that both the protection quality and the return potential of the dynamic VaR-strategy can be enhanced if a manager possesses superior forecast skills and if a high confidence level is chosen. Unfortunately, these suggestions are only of limited relevance for the investment practice. The main reason for the application of portfolio insurance strategies is to protect a portfolio against negative market developments in every market environment, independent of whether market declines can be forecasted or not. Moreover, if an investor possesses superior forecasting skills, other active (market timing) strategies may exist which generate higher returns compared to a protected passive stock market investment. Overall, our analysis contradicts the findings in Jiang et al. (2009). Based on our results, we conjecture that the main reason for the superiority of the dynamic VaR-strategy compared to the CPPI strategy in their study are the perfect quarterly return and volatility forecasts which they use to implement their strategy. Our simulation analysis suggests that their results are no longer robust when estimation risk is taken into account.

[17] This combined effect, however, is not observable for the synthetic put strategy. While a perfect volatility forecast also improves the protection quality of the synthetic put strategy, returns cannot be enhanced (see Tables II and VI).

5.5 Monte Carlo Simulations

5.5.1 Design of the Analysis

In this section, we check the robustness of our bootstrap results by additionally running Monte Carlo simulations. This approach allows us to analyze different economic scenarios with different stochastic parameters. The two main parameters that are required for the simulation of stock market returns are the expected return and its corresponding volatility. We define four states of nature to analyze the influence of these parameters on our results in a systematic way. In a first step, we distinguish between a high and a low equity risk premium state. According to Dimson et al. (2006), the long-run risk premium for developed stock markets is roughly 7% per year, this value represents our high risk premium state. They further argue that the equity risk premium might be lower in the years ahead and estimate the annual expected stock market excess return to be 4.5%. This conservative estimate represents our low risk premium state. Following Arnott and Bernstein (2002), we set the risk-free rate (money market rate) to a fixed value of 4.5% per year. Adding this value for the risk-free rate to the different equity risk premiums, we specify an expected stock market return of 9% in the low equity risk premium state and of 11.5% in the high equity risk premium state.

In the second step, we distinguish between a high and a low stock market volatility scenario. For the low volatility state, we assume a value of 20%, which roughly represents the average annual stock market volatility (Dimson et al. (2006)). In contrast, we use a stock market volatility of 30% per year for the high volatility state. Combining the two different values for the expected return (μ) and the volatility (σ), we have four different scenarios in our Monte Carlo simulations: scenario 1 ($\mu = 9\%; \sigma = 20\%$), scenario 2: ($\mu = 11.5\%; \sigma = 20\%$), scenario 3: ($\mu = 9\%; \sigma = 30\%$), and scenario 4: ($\mu = 11.5\%; \sigma = 30\%$). Based on these four scenarios, we simulate continuously compounded stock market returns on the basis of a geometric Brownian motion (Hull (2008)). This return-generating process has the main advantage of fulfilling the assumption of normally distributed returns underlying the synthetic put strategy and the dynamic VaR-strategy.

In order to derive the mean and volatility parameters for the geometric Brownian motion return-generating process, the expected return and volatility estimates are transformed into their corresponding continuously compounded counterparts. We simulate 250 daily stock market returns and implement the portfolio insurance strategies in the same way as discussed in Section 5.3.1 (i.e., without short sales and without leverage, 10 basis points round-trip transaction costs, a trading filter of 2%, and a protection level of 100%). Moreover, we use the scenario-based estimators for the equity risk premium and the volatility in implementing the dynamic VaR-strategy.

Given the stochastic character of the geometric Brownian motion, any generated return path may deviate from these 'true' parameters, also implying that we implicitly incorporate estimation risk. The necessary volatility for the synthetic put strategy is estimated in the same way. In order to derive the full distribution of the return, risk, and performance measures, we perform $10,000$ simulation runs.

5.5.2 Testing for Statistical Significance

The test procedure for deriving statistical significance is identical to the approach used in our block-bootstrap simulations (see Section 5.3.3). The only difference relates to the selection of elements from the return series generated from our block-bootstrap in the first step up until now. As the elements are now completely independent due to our Monte Carlo simulation set-up, we do not apply the stationary bootstrap of Politis and Romano (1994) in the second step, but instead use the standard bootstrap method suggested by Efron (1979) for all hypothesis tests. In particular, we draw individual elements rather than drawing blocks of elements. In the first step, $N = 10,000$ annual returns are generated by our Monte Carlo simulation. In order to derive confidence intervals, we then create $N_B = 5,000$ bootstrap resamples in the second step, each consisting of $10,000$ annual returns.

5.5.3 Monte Carlo Simulation Results

In order to validate the bootstrap results presented in Table V, we again use a multiplier of $m = 9$ for the CPPI strategy and the TIPP strategy in our Monte Carlo simulations, and the confidence level in the dynamic VaR-strategy is set to $\alpha = 0.95$. All Monte Carlo simulation results are presented in Table VII. In all four simulated market scenarios in Panel A of Table VII, the CPPI strategy dominates the stop-loss trading rule in terms of Omega ratios (comparison 1). The differences in Omega ratios are statistically significant in all scenarios and for all threshold levels (Panel B of Table VII). As in our bootstrap simulations, the main driver for this result is the superior downside protection quality of the CPPI strategy. Comparing the CPPI and the TIPP strategy (comparison 2), the Monte Carlo simulation results are also similar to those of our bootstrap simulations. While the TIPP strategy exhibits the lowest LPM values of all simulated protection strategies, its return potential is comparatively low. Moreover, the TIPP strategy dominates the CPPI strategy in terms of Omega ratios, but only at the cost of its lower returns. With an increasing threshold level τ, the difference in the Omega ratios decreases. For example, in the two scenarios with a high annual stock market return of 11.5%, the difference between the two Omega ratios is no longer statistically significant based on a threshold level of 4% (see Panel B of Table VII).

Table VII – Monte-Carlo Simulations

With respect to the measures of interest, this table shows the return, risk, and performance values of Monte-Carlo simulations as a robustness test. Based on four scenarios (with given return and volatility assumptions), continuously compounded stock market returns are simulated on the basis of a geometric Brownian motion. The expected return and volatility estimates are transformed into their continuously compounded counterparts. $m = 9$ denotes the multiplier of the CPPI and the TIPP strategy, and $\alpha = 0.95$ is the confidence level of the dynamic VaR-strategy. We simulate 250 daily stock market returns and implement the portfolio insurance strategies. The scenario-based estimators for the equity risk premium and the volatility are used in the implementation of the dynamic VaR-strategy. The required volatility for the synthetic put strategy is estimated in the same way. In order to derive the full distribution of the return, risk, and performance measures, we perform 10,000 simulation runs. Panel A presents the return, risk, and performance characteristics of the different portfolio insurance strategies. In addition to the mean annual returns, the strategies' protection quality is captured by using downside risk measures. Lower partial moments (LPMs) only take negative deviations from the target return (τ) into account. The target return is set equal to zero ($\tau = 0$), and we penalize negative deviations from this target with an exponent of 0 (LPM0) and 1 (LPM1). LPM0 and LPM1 represent the shortfall probability and the expected shortfall, respectively. The Omega measure, defined as the ratio of gains to losses relative to a target return, is used for our systematic comparison of the different portfolio insurance strategies along the distinguishing characteristics as discussed in Section 5.2.6 (comparisons 1 – 4 in Panel B). The threshold returns for the Omega ratio are set to 1%, 2%, 3%, and 4% per year, respectively. Panel B reports the confidence intervals for significance tests of the difference between the Omega ratios of each pair of strategies. In order to derive the confidence intervals, we implement a bootstrap using the $N = 10,000$ annual returns from the Monte-Carlo simulation in the first step. Within this second step bootstrap, we create $N_B = 5,000$ bootstrap resamples, each consisting of 10,000 annual returns. The standard bootstrap method suggested by Efron (1979) is used for all hypothesis tests in the second step. The confidence intervals in Panel B are evaluated at the 1%, 5%, and 10% level of significance. The confidence interval with the highest level of significance is reported. If the confidence interval based on a 10% significance level still contains zero, there are no significant differences in the Omega ratios of the respective pair of strategies. In contrast, if both reported interval boundaries are positive (negative), one strategy features a significantly higher (lower) Omega ratio than the other at the reported level of statistical significance. ***, **, and * denotes statistical significance at the 1%, 5%, and 10% level, respectively.

Panel A: Return, Risk, and Performance of Portfolio Insurance Strategies						
	Stop-Loss	Synthetic Put	CPPI ($m = 9$)	TIPP ($m = 9$)	Dynamic VaR ($\alpha = 0.95$)	Stock Market
Scenario 1: Return 9%, Volatility 20%						
Mean return p.a.	6.29%	6.29%	6.12%	5.03%	5.35%	9.29%
LPM0 ($\tau = 0\%$)	75.19%	31.48%	8.00%	1.36%	5.01%	34.31%
LPM1 ($\tau = 0\%$)	0.57%	0.32%	0.05%	0.01%	0.03%	3.75%
$\Omega(\tau = 1\%)$	5.00	8.35	15.53	57.53	22.64	3.02
$\Omega(\tau = 2\%)$	3.06	4.47	5.86	14.02	6.99	1.63
$\Omega(\tau = 3\%)$	2.16	3.81	3.16	4.98	3.26	2.29
$\Omega(\tau = 4\%)$	1.64	1.94	2.01	2.14	1.84	2.00
Scenario 2: Return 11.5%, Volatility 20%						
Mean return p.a.	7.69%	7.41%	7.14%	5.37%	6.21%	11.38%
LPM0 ($\tau = 0\%$)	71.06%	28.17%	7.91%	1.33%	5.30%	29.99%
LPM1 ($\tau = 0\%$)	0.53%	0.30%	0.06%	0.01%	0.03%	3.12%
$\Omega(\tau = 1\%)$	6.39	10.74	20.23	74.89	27.24	4.03
$\Omega(\tau = 2\%)$	3.91	5.81	7.78	18.08	8.81	3.50
$\Omega(\tau = 3\%)$	2.76	3.67	4.19	6.36	4.25	3.04
$\Omega(\tau = 4\%)$	2.09	2.54	2.65	2.72	2.45	2.65

continued

Table VII – *Continued*

Scenario 3: Return 9%, Volatility 30%						
Mean return p.a.	5.37%	5.63%	5.53%	4.75%	4.67%	8.75%
LPM0 (τ = 0%)	84.99%	48.72%	3.10%	0.35%	2.97%	43.27%
LPM1 (τ = 0%)	0.92%	0.32%	0.01%	0.00%	0.01%	7.36%
$\Omega(\tau = 1\%)$	3.46	6.07	8.82	35.77	17.07	1.99
$\Omega(\tau = 2\%)$	2.28	3.27	3.69	8.94	4.91	1.82
$\Omega(\tau = 3\%)$	1.68	2.13	2.21	3.49	2.32	1.66
$\Omega(\tau = 4\%)$	1.32	2.53	1.53	1.64	1.35	1.52
Scenario 4: Return 11.5%, Volatility 30%						
Mean return p.a.	6.69%	6.62%	6.52%	4.98%	5.28%	11.56%
LPM0 (τ = 0%)	82.26%	44.97%	2.82%	0.39%	2.56%	38.78%
LPM1 (τ = 0%)	0.86%	0.30%	0.01%	0.00%	0.01%	6.22%
$\Omega(\tau = 1\%)$	4.38	7.55	11.45	40.82	21.05	2.60
$\Omega(\tau = 2\%)$	2.87	4.07	4.71	10.33	6.07	2.36
$\Omega(\tau = 3\%)$	2.11	2.66	2.79	4.02	2.89	2.15
$\Omega(\tau = 4\%)$	1.65	1.91	1.92	1.89	1.70	1.96

The Monte Carlo simulation results for the comparison between the CPPI strategy and the synthetic put strategy deviate from those of the bootstrap simulations (comparison 3). Based on the Omega ratio, there is a dominance of the CPPI strategy. While the synthetic put strategy shows higher Omega ratios than the CPPI strategy in the bootstrap simulations for threshold returns of 3% and 4%, this effect no longer shows up in the Monte Carlo simulations. For the synthetic put strategy, normally distributed returns lead to lower LPM1 values, but neither to superior LPM0 values nor to higher returns. In contrast, compared to the bootstrap simulations, the LPM0 and LPM1 values of the CPPI strategy improve noticeably in a Monte Carlo framework. Therefore, the CPPI strategy seems to benefit more from all the assumptions underlying the geometric Brownian motion as the return generating process.

The results for the comparison of the synthetic put strategy and the dynamic VaR-strategy are similar for both the bootstrap and the Monte Carlo simulation (comparison 4). While the dynamic VaR-strategy dominates in terms of Omega ratios for threshold levels of 1% and 2% in our bootstrap simulations, the dynamic VaR-strategy also generates a significantly higher Omega ratio for a 3% threshold return in our Monte Carlo framework. Comparing the LPM0 and LPM1 values of the dynamic VaR-strategy in Table V with the corresponding values in Table VII, we observe superior risk properties for the dynamic VaR-strategy in our Monte Carlo simulations. In all scenarios, the shortfall probabilities (as measured by the LPM0) are at least as good as or even lower

Table VII – *Continued*

Panel B: Statistical Significance					
	Comparison 1	Comparison 2	Comparison 3	Comparison 4	
	CPPI versus Stop-Loss	CPPI versus TIPP	CPPI versus Synthetic Put	Dynamic VaR versus Synthetic Put	Dynamic VaR versus CPPI
Scenario 1: Return 9%, volatility 20%					
$\Omega(\tau = 1\%)$	[9.57;11.54] ***	[−47.85;−36.85] ***	[6.48;7.92] ***	[13.03;15.69] ***	[5.86;8.34] ***
$\Omega(\tau = 2\%)$	[2.51;3.09] ***	[−9.14;−7.28] ***	[1.22;1.57] ***	[2.25;2.79] ***	[0.86;1.41] ***
$\Omega(\tau = 3\%)$	[0.86;1.15] ***	[−2.08;−1.58]	[0.27;0.43] ***	[0.36;0.55] ***	[0.00;0.18] **
$\Omega(\tau = 3\%)$	[0.28;0.47] ***	[−0.23;−0.03] ***	[0.02;0.11] ***	[−0.15;−0.06] ***	[−0.24;−0.10] ***
Scenario 2: Return 11.5%, volatility 20%					
$\Omega(\tau = 1\%)$	[12.55;15.26] ***	[−63.88;−46.94] ***	[8.48;10.52] ***	[15.00;18.17] ***	[5.59;8.52] ***
$\Omega(\tau = 2\%)$	[3.49;4.26] ***	[−11.60;−9.10] ***	[1.74;2.20] ***	[2.70;3.33] ***	[0.71;1.34] ***
$\Omega(\tau = 3\%)$	[1.25;1.62] ***	[−2.50;−1.86] ***	[0.42;0.62] ***	[0.47;0.69] ***	[−0.03;0.14]
$\Omega(\tau = 3\%)$	[0.45;0.68] ***	[−0.15;0.02]	[0.06;0.17] ***	[−0.14;−0.03] ***	[−0.28;−0.12] ***

Table VII – *Continued*

Panel B: Statistical Significance					
	Comparison 1	Comparison 2	Comparison 3	Comparison 4	
	CPPI versus Stop-Loss	CPPI versus TIPP	CPPI versus Synthetic Put	Dynamic VaR versus Synthetic Put	Dynamic VaR versus CPPI
Scenario 3: Return 9%, volatility 30%					
$\Omega(\tau = 1\%)$	[4.81;5.91] ***	[−29.86;−24.56] ***	[2.37;3.16] ***	[10.08;11.98] ***	[7.34;9.16] ***
$\Omega(\tau = 2\%)$	[1.19;1.63] ***	[−5.76;−4.32] ***	[0.29;0.56] ***	[1.46;1.84] ***	[0.98;1.49] ***
$\Omega(\tau = 3\%)$	[0.40;0.66] ***	[−1.47;−1.11] ***	[0.02;0.14] **	[0.11;0.27] ***	[0.01;0.21] **
$\Omega(\tau = 3\%)$	[0.12;0.30] ***	[−0.21;−0.02] ***	[−0.04;0.03]	[−0.24;−0.13] ***	[−0.27;−0.09] ***
Scenario 4: Return 11.5%, volatility 30%					
$\Omega(\tau = 1\%)$	[6.37;7.76] ***	[−32.71;−26.40] ***	[3.43;4.38] ***	[12.40;14.70] ***	[8.57;10.70] ***
$\Omega(\tau = 2\%)$	[1.58;2.11] ***	[−6.26;−5.02] ***	[0.49;0.80] ***	[1.77;2.23] ***	[1.09;1.64] ***
$\Omega(\tau = 3\%)$	[0.53;0.84] ***	[−1.44;−1.02] ***	[0.05;0.22] ***	[0.15;0.33] ***	[0.01;0.19]
$\Omega(\tau = 3\%)$	[0.17;0.38] ***	[−0.05;0.10]	[−0.03;0.05]	[−0.27;−0.16] ***	[−0.32;−0.12] ***

than the expected values determined by the confidence level α. From this point of view, both the assumption of normally distributed returns and the negligence of deviating time series feature benefit the dynamic VaR-strategy more than the synthetic put strategy. The risk reduction effect of the dynamic VaR-strategy based on normally distributed returns may also provide an explanation why this strategy provides statistically significant higher Omega ratios than the CPPI strategy for all threshold returns ranging from 1% to up to 3% (depending on the scenario). This result differs from our bootstrap simulations, where no statistical significance of the dynamic VaR-strategy is observable against the CPPI strategy in terms of their Omega ratios for all threshold levels. Taken as a whole, however, our Monte Carlo simulation results confirm the bootstrap results for the comparison of the dynamic VaR-strategy with the synthetic put strategy as well as the CPPI strategy. Most importantly, our systematic comparison indicates that normally distributed returns are an important requirement for the protection quality of the dynamic VaR-strategy.

5.6 Conclusions

This study presents a systematic comparison of portfolio insurance strategies by using simulation analyses in order to conduct statistical significance tests. In addition to popular strategies, we also apply more recently developed protection strategies, such as the dynamic VaR-strategy. Our comparison of the different portfolio insurance strategies is based on the Omega ratio, which incorporates both the risk and return properties of a given strategy and does not depend on any distributional assumptions. In order to test the statistical significance of differences in Omega ratios, we use a bootstrap-based hypothesis test proposed by Politis and Romano (1994).

Our systematic comparison of different portfolio insurance strategies considers the following distinguishing characteristics: static vs. dynamic; initial wealth vs. cumulated wealth protection; model-based vs. model-free; and strong floor compliance vs. probabilistic floor compliance. Taken together, our findings reveal that the classical portfolio insurance strategies synthetic put and CPPI provide superior downside protection compared to the simple stop-loss trading rule, which ultimately results in significantly higher Omega ratios. Analyzing alternative portfolio insurance strategies, neither the TIPP strategy (as an 'improved' CPPI strategy) nor the dynamic VaR-strategy (with probabilistic floor compliance) are able to provide improvements over the classical strategies. A comparison of the dynamic VaR-strategy with the CPPI strategy and the synthetic put strategy shows no value added of this novel approach when estimation risk is incorporated into the analysis. While the CPPI strategy offers better downside protection than the dynamic VaR-strategy, the synthetic put strategy dominates the dynamic VaR-strategy in terms of its return potential. Most importantly, the usefulness of the latter strategy strongly depends on the forecasting quality for the equity risk premium and the stock market volatility as well as the assumption of normally distributed stock returns.

Once an investor has decided to use a protection strategy within his asset allocation strategy, our simulation results indicate an inverse relationship between the protection quality and the return potential of the different portfolio insurance strategies. As expected, strategies with a better downside protection (such as the CPPI strategy and the TIPP strategy with a low multiplier) tend to provide lower returns. Implementing the portfolio insurance strategy that best meets an investor's risk and return preferences, the superiority of a strategy when assessed by the Omega ratio depends on the level of the threshold return required by the investor. With higher threshold returns, strategies with a higher return potential become increasingly attractive. In most instances, the CPPI strategy exhibits a superior ratio of realized gains to losses compared to all other portfolio insurance strategies.

References

Annaert, J., S. Van Osselaer, and B. Verstraete (2009). "Performance Evaluation of Portfolio Insurance Strategies Using Stochastic Dominance Criteria." In: *Journal of Banking and Finance* 33, pp. 272–280.

Arnott, R.D. and P.L. Bernstein (2002). "What Risk Premium Is Normal?" In: *Financial Analysts Journal* 58, pp. 64–85.

Basak, S. (1995). "A General Equilibrium Model of Portfolio Insurance." In: *Review of Financial Studies* 8, pp. 1059–1090.

Benartzi, S. and R.H. Thaler (1995). "Myopic Loss Aversion and the Equity Premium Puzzle." In: *Quarterly Journal of Economics* 110, pp. 73–92.

Benninga, S. (1990). "Comparing Portfolio Insurance Strategies." In: *Financial Markets and Portfolio Management* 4, pp. 20–30.

Benninga, S. and M. Blume (1985). "On the Optimality of Portfolio Insurance." In: *Journal of Finance* 40, pp. 1341–1352.

Bertrand, P. and J.-L. Prigent (2005). "Portfolio Insurance Strategies: OBPI versus CPPI." In: *Finance* 26, pp. 5–32.

– (2011). "Omega Performance Measure and Portfolio Insurance." In: *Journal of Banking and Finance* 35, pp. 1811–1823.

Bird, R., R. Cunningham, D. Dennis, and M. Tippett (1990). "Portfolio Insurance: A Simulation under Different Market Conditions." In: *Insurance: Mathematics and Economics* 9, pp. 1–19.

Bird, R., D. Dennis, and M. Tippett (1988). "A Stop Loss Approach to Portfolio Insurance." In: *Journal of Portfolio Management* 15, pp. 35–40.

Black, F. and R. Jones (1987). "Simplifying Portfolio Insurance." In: *Journal of Portfolio Management* 14, pp. 48–51.

Black, F. and R. Litterman (1992). "Global Portfolio Optimization." In: *Financial Analysts Journal* 48, pp. 28–43.

Black, F. and M. Scholes (1973). "The Pricing of Options and Corporate Liabilities." In: *Journal of Political Economy* 81, pp. 637–654.

Bookstaber, R. and J.A. Langsam (2000). "Portfolio Insurance Trading Rules." In: *Journal of Futures Markets* 8, pp. 15–31.

Boyle, P.P. and T. Vorst (1992). "Option Replication in Discrete Time with Transaction Costs." In: *Journal of Finance* 47, pp. 271–293.

Brennan, M.J. and E.S. Schwartz (1989). "Portfolio Insurance and Financial Market Equilibrium." In: *Journal of Business* 62, pp. 455–472.

Brennan, M.J. and R. Solanki (1981). "Optimal Portfolio Insurance." In: *Journal of Financial and Quantitative Analysis* 16, pp. 279–300.

Cesari, R. and D. Cremonini (2003). "Benchmarking, Portfolio Insurance and Technical Analysis: A Monte Carlo Comparison of Dynamic Strategies of Asset Allocation." In: *Journal of Economic Dynamics and Control* 27, pp. 987–1011.

Choie, K.S. and E.J. Seff (1989). "TIPP: Insurance without Complexity: Comment." In: *Journal of Portfolio Management* 16, pp. 107–108.

Dichtl, H. and W. Drobetz (2011). "Portfolio Insurance and Prospect Theory Investors: Popularity and Optimal Design of Capital Protected Financial Products." In: *Journal of Banking and Finance* 35, pp. 1683–1697.

Dierkes, M., C. Erner, and S. Zeisberger (2010). "Investment Horizon and the Attractiveness of Investment Strategies: A Behavioral Approach." In: *Journal of Banking and Finance* 34, pp. 1032–1046.

Dimson, E., P. Marsh, and M. Staunton (2006). “Handbook of the Equity Risk Premium.” In: ed. by R. Mehra. Amsterdam: Elsevier. Chap. The Worldwide Equity Premium: A Smaller Puzzle.

Do, B.H. (2002). “Relative Performance of Dynamic Portfolio Insurance Strategies: Australian Evidence.” In: *Accounting and Finance* 42, pp. 279–296.

Do, B.H. and R.W. Faff (2004). “Do Futures-Based Strategies Enhance Dynamic Portfolio Insurance?” In: *Journal of Futures Markets* 24, pp. 591–608.

Dybvig, P. (1988). “Inefficient Dynamic Portfolio Strategies or How to Throw Away a Million Dollars in the Stock Market.” In: *Review of Financial Studies* 1, pp. 67–88.

Efron, B. (1979). “Bootstrap Methods: Another Look at the Jackknife.” In: *Annals of Statistics* 7, pp. 1–26.

Eling, M. and F. Schuhmacher (2007). “Does the Choice of Performance Measure Influence the Evaluation of Hedge Funds?” In: *Journal of Banking and Finance* 31, pp. 2632–2647.

Estep, T. and M. Kritzman (1988). “TIPP: Insurance without Complexity.” In: *Journal of Portfolio Management* 14, pp. 38–42.

Figlewski, S., N.K. Chidambaran, and S. Kaplan (1993). “Evaluating the Performance of the Protective Put Strategy.” In: *Financial Analysts Journal* 49, pp. 46–69.

Grossman, S. J. and Z. Zhou (1996). “Equilibrium Analysis of Portfolio Insurance.” In: *Journal of Finance* 51, pp. 1379–1403.

Harlow, W.V. and R.K.S. Rao (1989). “Asset Pricing in a Generalized Mean-Lower Partial Moment Framework: Theory and Evidence.” In: *Journal of Financial and Quantitative Analysis* 24, pp. 285–311.

Herold, U., R. Maurer, and N. Purschaker (2005). “Total Return Fixed-Income Portfolio Management.” In: *Journal of Portfolio Management* 31, pp. 32–43.

Herold, U., R. Maurer, M. Stamos, and H. Thanh Vo (2007). “Total Return Strategies for Multi-Asset Portfolios.” In: *Journal of Portfolio Management* 33, pp. 60–76.

Hull, J.C., ed. (2008). *Options, Futures, and Other Derivatives*. Prentice Hall International.

Jiang, C., Y. Ma, and Y. An (2009). “The Effectiveness of the VaR-based Portfolio Insurance Strategy: An Empirical Analysis.” In: *International Review of Financial Analysis* 18, pp. 185–197.

Kaplan, P.D. and J.A. Knowles (2004). "Kappa: A Generalized Downside Risk-Adjusted Performance Measure." In: *Journal of Performance Measurement* 8, pp. 42–54.

Künsch, H. R. (1989). "The Jacknife and the Bootstrap for General Stationary Observations." In: *Annals of Statistics* 17, pp. 1217–1241.

Leland, H.E. (1980). "Who Should Buy Portfolio Insurance?" In: *Journal of Finance* 35, pp. 581–594.

– (1985). "Option Pricing and Replication with Transactions Costs." In: *Journal of Finance* 40, pp. 1283–1301.

Liu, R. Y. and K. Singh (1992). *Moving Blocks Jacknife and Bootstrap Capture Weak Dependence*. Ed. by R. LePage and Y. Billard. New York: Wiley.

Patton, A., D.N. Politis, and H. White (2009). "Correction to "Automatic Block-Length Selection for the Dependent Bootstrap" by D. Politis and H. White." In: *Econometric Reviews* 28, pp. 372–375.

Politis, D.N. and J.P. Romano (1994). "The Stationary Bootstrap." In: *Journal of the American Statistical Association* 89, pp. 1303–1313.

Politis, D.N. and H. White (2004). "Automatic Block-Length Selection for the Dependent Bootstrap." In: *Econometric Reviews* 23, pp. 53–70.

Poterba, J.M. and L.H. Summers (1988). "Mean Reversion in Stock Prices: Evidence and Implications." In: *Journal of Financial Economics* 22, pp. 27–59.

Rendleman, R.J. and T.J. OBrien (1990). "The Effects of Volatility Misestimation on Option-Replication Portfolio Insurance." In: *Financial Analysts Journal* 46, pp. 61–70.

Rubinstein, M. and H.E. Leland (1981). "Replicating Options with Positions in Stock and Cash." In: *Financial Analysts Journal* 37, pp. 63–72.

Shadwick, W.F. and C. Keating (2002). "A Universal Performance Measure." In: *Journal of Performance Measurement* 6, pp. 59–84.

Zagst, R. and J. Kraus (2011). "Stochastic Dominance of Portfolio Insurance Strategies." In: *Annals of Operations Research* 185, pp. 75–103.

Zhu, Y. and R.C. Kavee (1988). "Performance of Portfolio Insurance Strategies." In: *Journal of Portfolio Management* 14, pp. 48–54.

Chapter 6

Dynamics of Time-Varying Volatility in the Dry Bulk and Tanker Freight Markets

with W. Drobetz and T. Richter

Acknowledgements

We thank two anonymous referees as well as Stefan Albertijn, Anna Merikas, and the participants of the 2010 annual meeting of the International Association of Maritime Economists (IAME) in Lisbon for their helpful comments.

Abstract

This study examines whether shocks from macroeconomic variables or asymmetric effects are more suitable for explaining the time-varying volatility in the dry bulk and tanker freight markets or whether both effects should be incorporated simultaneously. Using Baltic Exchange indices during the sample period from March 1999 to October 2011 on a daily basis, we separately analyze the impact of macroeconomic shocks and asymmetric effects on the conditional volatility of freight rates by using a GARCH-X model and an EGARCH model, respectively. Furthermore, we simultaneously investigate both effects by specifying an EGARCH-X model. Assuming not only a normal distribution but also a *t*-distribution in order to better capture the fat tails of error terms, three important conclusions emerge for modeling the conditional volatility of freight rates: (i) The assumption of a *t*-distribution is better suited than a normal distribution is. (ii) Macroeconomic factors should be incorporated into the conditional variance equation rather than into the conditional mean equation. In addition, the number of macroeconomic factors that exhibit explanatory power decreases under a *t*-distribution. (iii) While there seem to be no asymmetric effects in the dry bulk freight market, these effects are strongly pronounced in the tanker freight market. Our empirical findings have important implications for freight rate risk management.

6.1 Introduction

The volatility of freight rates is a driving force for investment and hedging decisions in the maritime industry. Compared to the volatility of financial assets, freight rate volatility is very high. For example, featuring an annual standard deviation of almost 40% from March 1999 to October 2011, the annual volatility of changes of the Baltic Capesize Index (BCI) was more than twice as high as the volatility of the Morgan Stanley Capital International (MSCI) World Stock Market Index. Moreover, the Baltic Panamax Index (BPI) fell by more than 95% from 11,425 to 440 index points between May 2008 and December 2008. Therefore, it is of utmost importance for ship owners, operators and shipping banks alike to have a thorough understanding of the properties of freight rate volatility in order to infer value-enhancing investment and risk management decisions.

The prior literature discusses three different approaches to derive volatility estimates: (i) historic volatility, (ii) implied volatility and (iii) volatility based on GARCH models. As several market upswings and downturns have shown, historic standard deviations could lead to inappropriate volatility estimates. Based on the assumption that all observations of the underlying sample period are equally weighted, the calculation of historic standard deviations implies that the latest developments on financial markets are not sufficiently taken into account. While the calculation of higher moments refers to specific time intervals and not to a particular point in time, investors are rather interested in a volatility measure that represents the current uncertainty perceived by the market. In order to derive such a 'point-in-time measure', a backward calculation of option pricing models offers an often-used alternative. Provided that the underlying model is correctly specified, the implied volatility assesses the markets' current risk perception. However, as option pricing models usually assume normally distributed returns, implied volatilities can be biased due to skewed and fat-tailed return distributions of the underlying asset. As shown in Section 6.4, this problem is especially severe for freight rates. If freight rate options need to be priced, this procedure cannot be used, either. Therefore, our study focuses on volatility estimates based on the class of Generalized Autoregressive Conditional Heteroscedasticity (GARCH) models.

Several earlier studies have already shown that GARCH models are suitable for capturing the observed volatility clustering of freight rates and vessel prices.[1] In order to better explain the time-varying volatility of freight rates, recent studies concentrate on two extensions of the GARCH specification originally proposed by Bollerslev (1986): (i) GARCH-X models that incorporate macroeconomic factors into the conditional variance equation, and (ii) asymmetric models that enable an analysis of potential asymmetric effects. For example, investigating the dynamics of the time-varying risk of monthly changes in second-hand tanker prices, Kavussanos (1996a) extends the standard GARCH model by using macroeconomic factors both in the conditional mean and in the conditional variance equation. The empirical results indicate that the oil price has a major impact on tanker volatility and should be incorporated as a macroeconomic factor into both the conditional mean and the conditional variance equation. In another empirical study, Kavussanos (1997) analyzes the dynamics of the conditional volatility of second-hand prices for dry bulk vessels. His findings reveal a positive relationship between the conditional volatility of Handysize as well as Panamax prices and interest rates and between the conditional volatility of Capesize prices and 1-year time charter rates.

In order to examine whether there is an asymmetric impact of past innovations on the time-varying volatility of dry bulk freights rates, Chen and Wang (2004) provide evidence that the impact of comparatively large innovations on the conditional variance is higher than the impact of smaller innovations. Correcting for this asymmetric size effect by using the Exponential GARCH (EGARCH) model of Nelson (1991), they also document that negative changes in freight rates have a higher influence on the freight rate volatility than positive changes do. A similar analysis of Hui et al. (2008)

[1] For example, see Kavussanos (1996a, 1996b, 1997, 2003), Glen and Martin (1998), Alizadeh and Kavussanos (2002), Chen and Wang (2004), Hui et al. (2008), Alizadeh, Nomikos, and Dellen (2011), Alizadeh and Nomikos (2009, 2011), and Roumpis and Syriopoulos (2009).

confirms the evidence for an asymmetric size effect in the dry bulk shipping market. However, adjusting for the different magnitude of shocks, they report mixed results with respect to whether there are asymmetric effects and – if so – whether these effects are positive or negative. Roumpis and Syriopoulos (2009) use alternative dynamic volatility models to investigate the risk-return characteristics of shipping stocks. Their results indicate that the average impact of negative returns of shipping stocks on return volatility is stronger than that of equivalent positive shocks. Methodologically, the empirical results in Roumpis and Syriopoulos (2009) indicate that the EGARCH specification provides superior results compared to Ding et al.'s (1993) Asymmetric Power GARCH (APGARCH); the results of the latter model are not always robust. Most recently, Alizadeh and Nomikos (2011) examine the relationship between the dynamics of the term structure of freight rates and the conditional volatility of freight rates in both the dry bulk and the tanker freight market.[2] Estimating an EGARCH-X model that incorporates both macroeconomic factors and asymmetric effects, they document that the term structure of freight rates exhibits explanatory power with regard to the time-varying risk of freight rates. Moreover, they provide evidence for asymmetrical effects both in the bulker and tanker segment.

Overall, the previous academic literature reports that both macroeconomic factors and asymmetric effects are able to capture the time-varying risk of freight rates, helping to improve freight rate risk management. The more appropriate the volatility estimates are, the better freight rate risks can be managed. Therefore, the question of interest for our empirical analysis is whether shocks from macroeconomic factors or asymmetric effects are better suited for modeling the conditional volatility of freight rates. It could be even the case that both forces are necessary to capture freight rate heteroscedasticity and hence should be used simultaneously.

In this article, we therefore investigate both issues – a potential asymmetric impact of shocks and the influence of macroeconomic factors on time-varying volatility – separately and simultaneously for the dry bulk and the tanker freight markets. Methodologically, we work with a GARCH-X model, an EGARCH model as well as an EGARCH-X model and compare their ability to explain volatility clustering. In contrast to earlier studies, the daily frequency of our underlying Baltic Exchange indices and the long sample period from March 1999 to October 2011 provides the large number of observations that is necessary in order to derive robust volatility estimates. Finally, despite extensive prior research, error distributions other than the normal distribution have not been examined for freight rates so far. In order to better capture fat tails of the corresponding error distribution, we therefore assume not only a normal distribution but also a *t*-distribution. Analyzing potential model specifications for freight rate volatility,

[2] The term structure of freight rates is defined as the difference between long- and short-term period freight rates and is used as a proxy for backwardation and contango.

our article makes three major contributions: (i) The assumption of a t-distribution is better suited than a normal distribution is. (ii) Macroeconomic factors should be included in the conditional variance equation, but not in the conditional mean equation. Furthermore, the number of macroeconomic factors that exhibit explanatory power decreases under the t-distribution. (iii) While there seem to be no asymmetric effects in the dry bulk freight market, these effects are highly pronounced in the tanker freight market.

The remainder of this article is structured as follows: Section 6.2 discusses the theoretical foundations to apply GARCH-X and EGARCH models for freight rates. Section 6.3 provides a brief methodological introduction, while Section 6.4 describes the data and the descriptive statistics. Section 6.5 discusses our empirical findings. The article concludes with Section 6.6 and points out possible implications for market participants.

6.2 Theoretical Foundations

In order to detect potential driving forces of freight rate volatility, we concentrate on macroeconomic factors and on asymmetric effects. The prior literature provides evidence that both approaches are able to capture the time-variation of freight rate risk. Therefore, in a first step, we substantiate the questions why and which macroeconomic factors potentially contribute to explaining the conditional volatility of freight rates. In a second step, we lay the theoretical foundations for the presence of asymmetric effects in freight rate volatility.

As in any other market, freight rates in the shipping industry are formed by the interaction of supply and demand for sea transport. Therefore, it is necessary to analyze whether macroeconomic factors have an impact on the demand and supply for shipping services. The demand for sea transport arises from the need of exporters and importers to transport freight to specific destinations around the world.[3] This 'derived' demand is mainly affected by the global economy and by global trade, supporting the notion that macroeconomic variables, such as the Gross Domestic Product (GDP), the industrial production, or steel production, incorporate substantial information about the demand for shipping services. The better the condition of the global economy is, the higher global trade and therefore the need for sea transport will be. Moreover, random shocks based on political events or financial crises as well as cyclical and seasonal market movements of the commodities transported by sea (such as oil, iron ore, coal, metal, or wheat) further substantiate that the demand for sea transport depends on macroeconomic factors. Besides these macroeconomic factors, the distance between

[3] The demand for sea transport is measured in terms of tonne-miles which are defined as the tonnage of freight transported by sea multiplied by the average distance over which the cargo is shipped.

the places of production and consumption of products as well as transportation costs constitute other demand determinants.

Turning to the supply side, sea transport is mainly determined by the size of the world fleet measured in tonne-miles. Other important factors are the fleet productivity, the shipbuilding production, scrapping and losses, and the level of freight rates. The higher the freight rates are, the higher the incentive for ship owners will be to extend supply if at all possible. Taken together, according to Stopford's (2009) shipping market model, the key determinants of the demand for sea transport are the global economy, seaborne commodity trades, random shocks due to political or macroeconomic events, average haul distance, and transportation costs, whereas the supply for shipping services is influenced by the world fleet, fleet productivity, shipbuilding production, scrapping and losses, and freight revenues. Table I summarizes the impact on demand and supply of an increase of these variables. All in all, as freight rates are determined by the interaction of supply and demand for sea transport, we expect that macroeconomic factors will influence both the level of freight rates and their volatility.

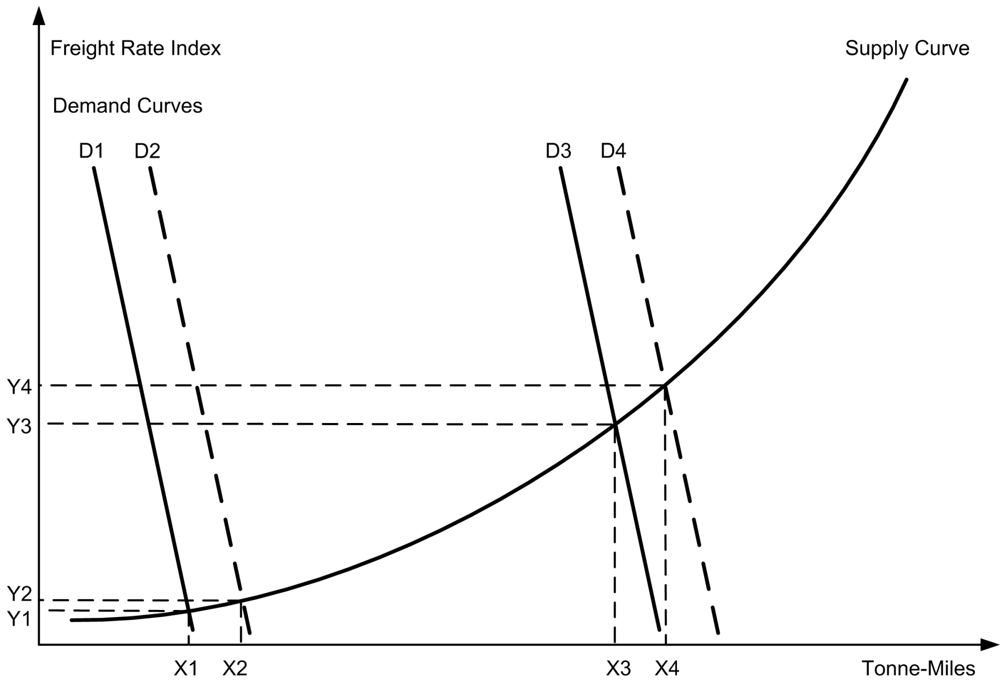
Table I – Determinants of the Shipping Market

Based on Stopford (2009), this table presents the determinants of the shipping market. Random shocks can have a negative or positive impact on the demand for sea transport, depending on whether a shock is caused by bad or good news.

Demand			Supply		
1	World Economy	+	1	World Fleet	+
2	Seaborne Commodity Trades	+	2	Fleet Productivity	+
3	Average Haul Distance	+	3	Shipbuilding Production	+
4	Random Shocks	−/+	4	Scrapping and Losses	−
5	Transportation Costs	+	5	Freight Revenues	+

The theoretical foundation for the presence of asymmetric effects of freight rate changes can also be inferred from a simple demand and supply model. Determined by the intersection of the supply and demand curves for shipping services, the equilibrium freight rate ensures that the quantity demanded by charterers is equal to the quantity supplied by ship owners. Figure I illustrates that the demand curve for shipping services is characterized by a comparatively sharp downward slope to the right. The lower the freight rate is, the higher the quantity demanded is. The demand curve is highly inelastic based on the fact that there is no competing mode of transportation in terms of costs. This notion is further strengthened by the low value of the freight cost in relation to the final price of the goods transported by sea. In contrast to the inelastic demand curve, the supply curve for shipping services is convex. It is highly elastic at low freight rate levels and becomes almost perfectly inelastic at high freight rate levels. When the global economy is in recession, freight rates are low. Accordingly, ships at sea reduce their speed, and some ships are laid up if the freight rate falls below their operating costs, which strongly depend on the vessel's technological state. Old and inefficient ships may even be driven out of the market for demolition. At this point, an increase in the demand of sea transport from X1 to X2 in Figure I only leads to a small increase in freight rates from Y1 to Y2. However, as the global economy recovers, freight rates increase and vessels start to speed up again. At the same time and in consideration of their operating costs, more and more ships are taken out of lay and added to the active fleet. After all ships have been reactivated, a further increase in demand for sea transport from X3 to X4 leads to a substantially higher increase in freight rates from Y3 to Y4 as shown in Figure I. This extra demand of sea transport can be only partially absorbed by operating the ships at maximum speed in order to carry as much freight as possible.

Figure I – The Shipping Market and Demand Model
Source: Stopford (2009)



Due to the time delay between placement of order and delivery of the ship ('time-to-build lag'), the supply curve becomes almost inelastic in the short-run. Only in the long-run, new ships can be built and added to the existing fleet, shifting the entire supply curve to the right. Overall, Stopford's (2009) shipping market supply and demand model predicts larger shocks in upswing markets and smaller shocks in recessions. In addition to this asymmetric size effect, the model posits that positive shocks have a higher impact on the conditional volatility than negative shocks of the same magnitude do based on the convexity of the supply curve.

6.3 Empirical Methodology

Engle (1982) introduces the class of Autoregressive Conditional Heteroscedasticity (ARCH) models, which assumes that the conditional variance is a function of lagged squared residuals. Bollerslev (1986) extends these types of models to the class of GARCH models, additionally incorporating lagged terms of the conditional variance. Being more parsimonious than ARCH models, GARCH models lead to more robust conditional variance estimates. All classes of GARCH models are characterized by the following three constituents: (i) the conditional mean equation, (ii) the conditional variance equation, and (iii) the density function of the innovation process. In our study, we use an autoregressive model with macroeconomic factors as the conditional mean equation:

$$\Delta fr_t = c + \sum_{m=1}^{a} \gamma_m \Delta fr_{t-m} + \sum_{n=1}^{b} \varphi_n X_n(l) + \epsilon_t, \tag{1}$$

where Δfr_t denotes the change of the underlying Baltic Exchange index from period $t-1$ to t, a the order of the autoregressive terms, b the number of macroeconomic factors included, l the time lag of the n-th macroeconomic factor X_n, and ϵ_t the corresponding innovation (error term) at time t. The conditional variance equation in Bollerslev's (1986) standard GARCH model is expressed as follows:

$$\sigma_t^2 = \omega + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2, \tag{2}$$

where p specifies the order of past conditional variances (GARCH terms) and q the order of past squared residuals (ARCH terms). Furthermore, the model must satisfy the following stationarity and non-negativity restrictions:

$$\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j < 1; \quad \omega > 0; \quad \alpha_i \geq 0; \quad \beta_j \geq 0 \quad \forall\, i \in \{1,\ldots,q\} \wedge j \in \{1,\ldots,p\}. \tag{3}$$

The first parameter constraint ensures the existence of a finite, time-independent variance of the innovation process. In order to assure a strictly positive conditional variance, additional non-negativity constraints are necessary.[4] Building on Engle's (1982) and Bollerslev's (1986) analyzes, several extensions of the standard models have

[4] These stationarity and non-negativity constraints of univariate GARCH models can be substantially weakened for $p > 2$ (Cao and Nelson (1992)).

been developed. Based on the theoretical foundations in Section 6.2, our study focuses on two classes of GARCH models: (i) GARCH-X models, which are characterized by incorporating macroeconomic variables into the conditional variance equation, and (ii) EGARCH models, which enable investigating possible asymmetric effects.

In a first step, we examine both effects separately. Analyzing the influence of macroeconomic factors on the conditional volatility of freight rates, we apply the GARCH-X model:

$$\sigma_t^2 = \omega + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{k=1}^{x} \theta_k X_k(l), \tag{4}$$

where $X_k(l)$ denotes the k-th macroeconomic variable with a time lag of the order l. If macroeconomic variables exhibit information content for explaining the time-varying conditional variance, the estimated parameter θ_k should be significant.

Depending on the information content of past innovations, the Exponential GARCH (EGARCH) model proposed by Nelson (1991) incorporates potential asymmetric effects of unanticipated shocks into the conditional variance equation:

$$\sigma_t^2 = \exp\left(\omega + \sum_{j=1}^{p} \beta_j \ln(\sigma_{t-j}^2) + \sum_{i=1}^{q} \delta_i \left| \frac{\epsilon_{t-i}}{\sigma_{t-i}} \right| + \sum_{i=1}^{q} \alpha_i \frac{\epsilon_{t-i}}{\sigma_{t-i}} \right). \tag{5}$$

In contrast to other asymmetric GARCH models, such as the GJR-GARCH model (Glosten et al. (1993)) or the Threshold GARCH (TGARCH) model (Zakoian (1994)), the EGARCH model accounts not only for the sign of shocks but also for potential differences in the magnitude of asymmetric shocks. Presumably, the impact of large unanticipated shocks will have a larger effect on the conditional variance than the impact of small unanticipated shocks. A positive (negative) value of δ_i indicates a size effect regarding the asymmetric impact of unanticipated shocks on the conditional variance of freight rates: the impact of comparatively large (small) innovations on the conditional variance is higher compared to the impact of comparatively small (large) innovations. Furthermore, controlling for this asymmetric size effect and based on the shipping supply and demand model, a positive (negative) value of α_i suggests that positive (negative) innovations exert a stronger impact on the conditional variance than negative (positive) innovations of the same magnitude do. Based on the exponential function, estimation problems relating to the non-negativity constraints no longer emerge. In a second step, we simultaneously analyze both effects, i.e., macroeconomic variables and asymmetric shocks, by using the following EGARCH-X model:

$$\sigma_t^2 = \exp\left(\omega + \sum_{j=1}^{p} \beta_j \ln(\sigma_{t-j}^2) + \sum_{i=1}^{q} \delta_i \left|\frac{\epsilon_{t-i}}{\sigma_{t-i}}\right| + \sum_{i=1}^{q} \alpha_i \frac{\epsilon_{t-i}}{\sigma_{t-i}} + \sum_{k=1}^{x} \theta_k X_k(l)\right). \tag{6}$$

Including both macroeconomic factors and terms that account for asymmetric effects in the conditional variance equation, this EGARCH-X model allows a simultaneous analysis of both driving forces. Specifically, it can help to determine whether shocks from macroeconomic variables or asymmetric effects are better suited for explaining the time-varying volatility in the dry bulk and tanker freight markets. It could be even the case that both forces have a significant explanatory power and should therefore be included in the conditional variance equation in order to derive appropriate volatility estimates for freight rates. The coefficients in equation (6) follow the same interpretation as in equations (4) and (5). Finally, with regard to the choice of the error term in (1), ϵ_t, we not only assume normally distributed but also model t-distributed error terms in order to better capture fat tails.

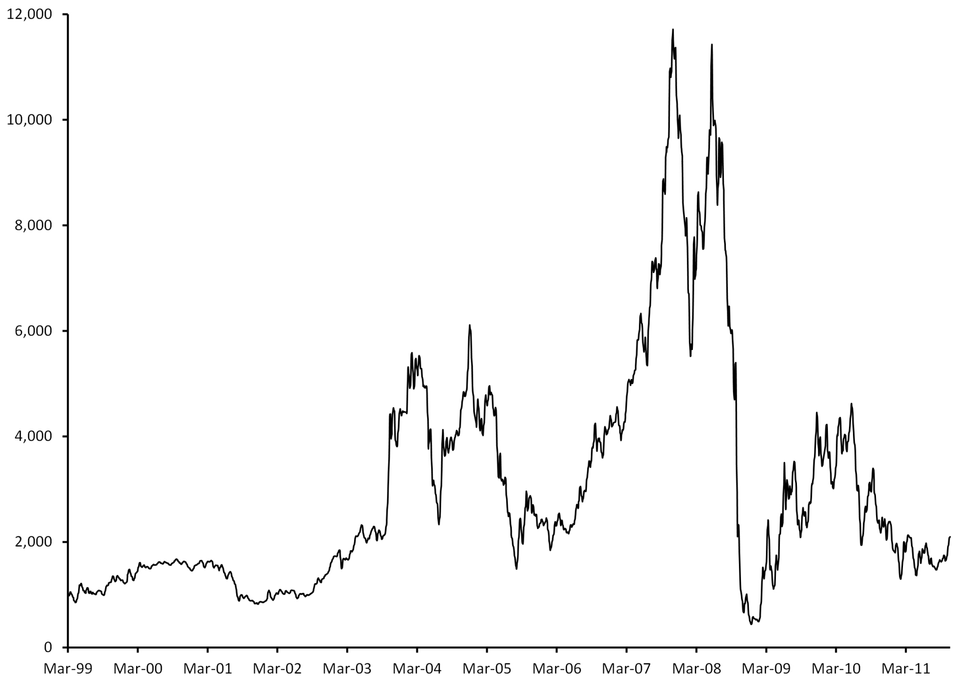
6.4 Data and Descriptive Statistics

Incorporating macroeconomic factors and asymmetric effects into the conditional variance equation requires a sufficiently high number of observations in order to derive robust volatility estimates. Being available on a daily basis since 1 March 1999 at the latest, the Baltic Exchange indices of the dry bulk and the tanker freight market fulfill this necessary requirement.[5] Depending on the underlying vessel type, each of these freight rate indices is calculated as a weighted average of freight rates on major voyage and time charter routes and is then converted into an index number that refers to a pre-specified basis. The freight rate of every single route is derived from the Baltic Exchange as the weighted arithmetic mean of a panel of independent shipbrokers. Our empirical work focuses on the dry bulk and the tanker freight market of the maritime industry, covering a sample period from 1 March 1999 to 18 October 2011. In particular, we analyze the conditional volatility of daily changes of the following four Baltic Exchange indices provided by Clarksons Shipping Intelligence Network: the Baltic Panamax Index (BPI), the Baltic Capesize Index (CPI), the Baltic Clean Tanker Index (BCTI), and the Baltic Dirty Tanker Index (BDTI).

Economically, our choice of indices is also justified by the fact that Baltic Exchange indices are used as underlying of freight options contracts. Since volatility is a major

[5] The container freight market has to be excluded from our analysis because there is no daily data available for our period under investigation.

Figure II – Baltic Panamax Index
Source: Clarksons Shipping Intelligence Network.



parameter for the pricing of freight options, more accurate and transparent volatility estimates will contribute to the development of the still illiquid freight option market. Moreover, the maritime industry's need for a well-functioning risk management is illustrated in Figure II. As an example, it depicts the BPI over the sample period from 1 March 1999 to 18 October 2011, reflecting the extreme risks in the shipping industry. Freight rates have achieved levels in height never seen before in the shipping industry since mid-2003. In particular, this is valid for October to December 2007 and for May 2008. Due to the fact that the demand for shipping services is strongly related to the global economy, the outbreak of the financial crisis in mid-2008 has led to a sharp decrease in freight rates. For instance, the BPI dropped by more than 95% within a few months. Therefore, the maritime industry must be considered as extremely volatile. This notion is further supported by the descriptive statistics of the indices presented in Table II. For example, the extremely high standard deviations in all shipping subsectors and the extreme market movements as high as 38.1% and 25.4% per day characterize the shipping industry as one of the riskiest industries worldwide. The very low p-values of the Jarque-Bera (1980) test confirm that Baltic Exchange indices are neither daily log-normally distributed (results not shown), nor are their daily changes normally distributed. Most important, the probability of fat tails is much higher than predicted by a normal distribution.[6]

[6] This finding has important implications regarding the calculation of implicit volatilities. Option

Table II – Descriptive Statistics of Daily Changes of Baltic Exchange Indices
Provided by Clarksons Shipping Intelligence Network, this table presents Baltic Exchange freight rate indices, where BPI denotes the Baltic Panamax Index, BCI the Baltic Capesize Index, BDTI the Baltic Dirty Tanker Index, and BCTI the Baltic Clean Tanker Index. The calculations are based on logarithmic differences of the indices. Jarque-Bera denotes the p-value of the Bera and Jarque (1980) test for normality. ARCH (3) describes the p-value of the Lagrange multiplier test for 3rd order autoregressive conditional heteroscedasticity in the residuals (Engle (1982)).

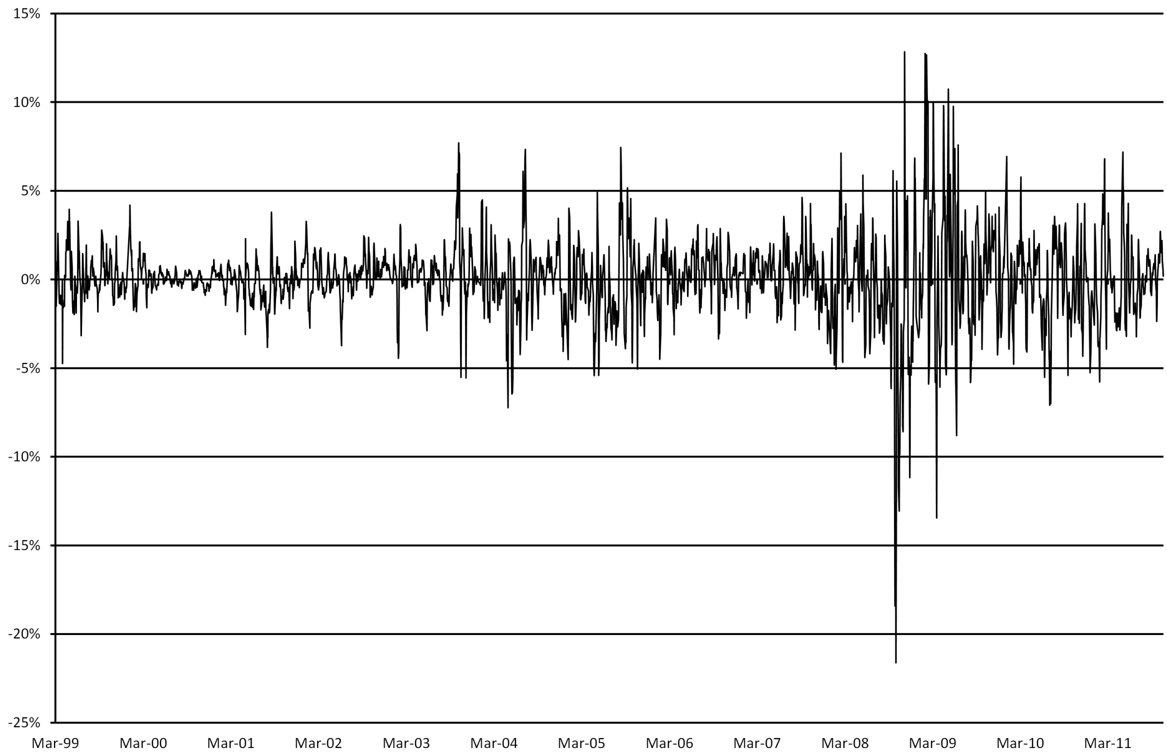
Statistics	BPI	BCI	BDTI	BCTI
Annualized Mean	0.054	0.094	−0.005	0.007
Annualized Volatility	0.349	0.399	0.339	0.225
Skewness	−0.526	0.117	−1.555	−1.035
Kurtosis	13.403	10.296	43.226	99.511
Minimum	−0.216	−0.192	−0.381	−0.296
Maximum	0.128	0.165	0.230	0.254
Jarque-Bera (p-value)	0.000	0.000	0.000	0.000
ARCH (3) (p-value)	0.000	0.000	0.000	0.000

Not only are freight rates extremely volatile, but their risk is also time-varying. As daily changes of the BPI exhibit low as well as high volatility cycles, Figure III suggests that the assumption of homoscedasticity cannot be maintained. Another observation is that daily changes of the BPI increased over time during the sample period. In order to confirm our observations statistically, we test for the presence of ARCH effects in the residuals by applying Engle's (1982) Lagrange multiplier test. Analyzing several autoregressive models for all four Baltic Exchange indices, the null hypothesis of no ARCH effects must be rejected in all cases, regardless of the number of time lags in the residuals. Reporting a p-value of below 0.01 in all cases, heteroscedasticity seems to be highly pronounced in all indices.

Our empirical analysis which macroeconomic factors potentially influence the conditional volatility of freight rates builds on Stopford's (2009) shipping market model, suggesting that there are ten determinants that affect the supply and demand for sea transport and hence the level as well as the volatility of freight rates (see Table I). However, as our examination requires daily data in order to derive robust coefficient estimates, most of these determinants cannot be included in the conditional variance equation because they are only available on a monthly basis, if at all. In particular, these are the average haul distance, the size of the world fleet and its productivity (e.g., laid-up tonnage), shipbuilding production, and scrapping and losses. For the same reason, other macroeconomic variables such as the GDP, the industrial production, and the inflation rate must be excluded from our analysis as well. The remaining determinants – the global economy, seaborne commodity trades, random shocks and

pricing formulas that require a normal or a lognormal distribution (for example, Modified Black (1976), Turnbull-Wakeman approximation (1991), Levy (1997), or Koekebakker et al. (2007)) should not be applied. Otherwise, estimation biases could occur, clearly showing the need for distributions that are able to capture fat tails more appropriately.

Figure III – Daily Changes of the Baltic Panamax Index
Source: Clarksons Shipping Intelligence Network.



transportation costs – all meet the necessary requirement of data availability on a daily basis although they cannot be measured directly. Therefore, we use the following proxy variables that are able to capture the information content of these determinants on a daily basis: the MSCI World Index as a global stock market index, the London Brent Crude Oil Index, the Kansas Hard Wheat Index as a proxy variable for wheat prices, the London Metal Exchange Index, the Goldman Sachs Commodity Index (GSCI), the Treasury-Eurodollar (TED) spread, and the spread of the yield curve.[7]

Table III presents our hypotheses concerning the magnitude of the impact of these macroeconomic factors on the conditional volatility of freight rates. As the demand for sea transport is a derived demand, one would expect a positive relationship between the global stock market, oil, wheat, and metal prices as well as the GSCI and the determinants global economy and seaborne commodity trades. Presumably, the higher (lower) the levels of these proxy variables are, the better (worse) is the state of the global economy and the higher (lower) are seaborne commodity trades, and vice versa. Furthermore, we assume that a high volatility of these macroeconomic variables will also lead to a high volatility of freight rates. Large changes of the MSCI World Index signal a high uncertainty of the market participants, and hence we also anticipate a high

[7] Appendix A provides detailed information about these macroeconomic factors.

Table III – Impact of Macroeconomic Factors on Freight Rate Volatility
This table illustrates the impact of macroeconomic factors on freight rate volatility. While ++ denotes the expectation of a very strong impact of the corresponding proxy variable on the conditional volatility of freight rates, + quantifies a strong effect. If a cell is blank, the respective macroeconomic factor is assumed to have no or only a limited influence on the conditional volatility of freight rates.

Macroeconomic Factor	Global Economy	Seaborne Commodity Trades	Random Shocks	Transportation Costs
Stock Market	+	+	++	+
Oil Price	+	++	++	++
Wheat Price	+	++	++	
Metal	+	++	++	
Commodity Price	+	++	++	+
TED Spread	+	+	++	
Term Spread	++	+	++	

volatility of freight rates. However, we expect that the impact of these proxy variables on the volatility of freight rates varies in terms of their explanatory power. For example, as indicated by the number of plus signs in Table III, we assume a highly pronounced impact of the global stock market index on the global economy and a strong influence on seaborne commodity trades. With regard to the commodity indices, we expect a similar pattern. The oil, wheat, and metal indices as well as the GSCI should all have a very strong impact on the seaborne commodity trades and a strong effect on the global economy. Another measure for market participants' uncertainty is the TED spread (Ferson and Harvey (1994)), which is defined as the difference between the 3-month money market rate on interbank loans and the yield for 90-day US Treasury Bills. A higher TED spread is an indicator of lower current investor sentiment and higher credit risk, i.e. an increasing (decreasing) TED spread indicates higher (lower) overall risk in the global financial system. Accordingly, a high (low) TED spread should also lead to a high (low) conditional volatility of freight rates. Being a reliable predictor of the future real economic activity, we also consider the spread (slope) of the yield curve, defined as the difference between 10-year US Treasury Notes and 3-month US Treasury Bills (Estrella and Hardouvelis (1991), Harvey (1991)). A high (low) spread of the yield curve indicates a good (bad) state of the economy. Competing with the stock market, long-term interest rates must be high in upswing markets in order to induce market participants in their intertemporal investment-consumption decision to invest into government bonds. Given that Stopford's (2009) shipping market supply and demand model predicts comparatively small shocks in recessions, but larger shocks in upswing markets, we expect a positive relationship between the slope of the yield curve and the conditional volatility of freight rates. In addition, as random shocks can be caused by political events, wars, or even the weather, we assume that the resulting market distortions are already captured to a great extent by our suggested proxy variables. Finally, we anticipate that transportation costs are mainly driven by the oil price. Taken together, our expectations are confirmed by the descriptive statistics of the macroeconomic factors in Table IV. The high kurtosis, the minimum and maximum changes, and the significant deviation from the normal distribution all indicate that large changes of each of these macroeconomic variables tend to occur frequently.

Table IV – Descriptive Statistics of Daily Changes of Macroeconomic Variables

Provided by Thomson Reuters Datastream, this table presents the descriptive statistics of daily changes of macroeconomic variables. In particular, the following indices are used: the MSCI World Index as a global stock market index, the London Brent Crude Oil Index, the Kansas Hard Wheat Index as a proxy variable for wheat prices, the London Metal Exchange Index, the Goldman Sachs Commodity Index (GSCI), the Treasury-Eurodollar (TED) spread, and the spread of the yield curve (term spread). The calculations of the macroeconomic factors stocks, oil, metal, wheat and GSCI are based on logarithmic differences. The calculations of the term spread and the TED spread are based on levels. Jarque-Bera denotes the p-value of the Jarque and Bera (1980) test for normality. ARCH (3) describes the p-value of the Lagrange multiplier test for 3rd order autoregressive conditional heteroscedasticity in the residuals (Engle (1982)).

Statistics	Stocks	Oil	Metal	Wheat	GSCI	TED Spread	Term Spread
Annualized Mean (%)	0.30	17.63	9.23	6.75	7.15	1.65	1.89
Annualized Volatility (%)	17.37	29.10	24.37	31.99	24.69	19.91	20.28
Minimum (%)	−7.33	−11.35	−9.39	−12.30	−9.17	−0.24	−0.60
Maximum (%)	9.10	9.77	8.23	12.54	7.22	5.72	3.79
Skewness	−0.34	−0.43	−0.36	0.05	−0.31	0.67	−0.31
Kurtosis	10.26	5.59	6.33	6.40	5.39	2.62	1.73
Jarque-Bera (p-value)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ARCH (3) (p-value)	0.00	0.01	0.06	0.00	0.65	0.00	0.66

The correlation matrix of the macroeconomic factors as well as the Variance Inflation Factors (VIFs) are shown in Table V.[8] The TED spread and the yield curve exhibit the highest correlation with a value of 0.44, indicating that these factors could be driven by the same latent variable. We test whether multi-collinearity is a potential problem in our volatility specifications by calculating VIF-values. As indicated in Table V, all values are below 10, suggesting that potential estimation biases due to multi-collinearity are negligible.

[8] The VIF is a measure of how much the variance of the estimated regression coefficient is 'inflated' by the existence of correlation among the regressors in the model. If significant multi-collinearity exist, the VIF will be very large ($VIF > 10$) for the corresponding variable. Table V shows the results of uncentered VIFs. The findings for centered VIFs are even better with no value higher than 1.36.

Table V – Descriptive Statistics of Daily Changes of Macroeconomic Variables
This table reports the correlation matrix of the macroeconomic factors used in this study. All data is provided by Thomson Reuters Datastream. Correlations are based on logarithmic differences. The last column shows the corresponding Variance Inflation Factors (VIF).

	Stocks	Oil	Metal	Wheat	GSCI	TED Spread	Term Spread	VIF
Stocks	1.00	−0.00	0.40	0.15	0.31	−0.04	−0.01	1.22
Oil		1.00	−0.01	−0.04	−0.01	−0.04	−0.01	1.00
Metal			1.00	0.20	0.42	−0.05	−0.01	1.36
Wheat				1.00	0.28	−0.03	−0.02	1.10
GSCI					1.00	−0.03	−0.01	1.31
TED Spread						1.00	0.44	1.40
Term Spread							1.00	1.97

6.5 Empirical Results

Prior empirical studies and our own preliminary analysis document that freight rates do not exhibit constant volatility. Based on the theoretical foundations presented in Section 6.2, we apply the GARCH-X, the EGARCH, and the EGARCH-X model in order to analyze whether shocks from macroeconomic variables or asymmetric effects are more suitable for modeling the time-variation of freight rate risk, or whether both driving forces should be considered simultaneously. Our criterion for model selection is the Akaike Information Criterion (AIC) proposed by Akaike (1974). Specifically, the AIC captures the trade-off between the statistical goodness of fit and the complexity of the model by imposing a penalty for increasing the number of parameters that have to be estimated:

$$AIC = -2\frac{LL}{n} + \frac{2k}{n}, \tag{7}$$

where LL denotes the maximized log-likelihood of the model, k the number of parameters, and n the number of observations. Comparing any two model results, the AIC recommends the specification with the lower value, implying either an improved model fit or fewer explanatory variables, or both.

6.5.1 GARCH-X Model

Our empirical analysis of conditional volatility starts by extending the standard GARCH(p,q) framework to the GARCH-X model. In order to examine whether macroeconomic variables capture any substantial information for explaining the conditional variance, we include them (i) in the conditional mean equation, (ii) in the conditional variance equation, and (iii) in both the conditional mean and the conditional variance equation. Based on the model selection according to the AIC, Table VI reports our empirical findings.

The conditional mean equation reveals that all daily freight rates exhibit strongly pronounced autocorrelation, as indicated by the autoregressive terms that are highly significant up to the order of 3.[9] Therefore, a large fraction of today's changes in freight rates can be attributed to changes in freight rates of the last two or three trading days, respectively. Most important, the high coefficient γ_1 of the first autoregressive term accounts for a large portion of the highly pronounced heteroscedasticity of freight rate changes. This high autocorrelation can be explained by the limited substitution opportunities for shipping services over time as well as substitution across routes and vessel types (Alizadeh et al. (2007), Bessler et al. (2008)). In order to ensure the highest possible degree of capacity utilization, ship owners are interested in immediate follow-on employments of their vessels, and hence they can neither speculate on increasing freight rates by delaying the vessel's loading, nor can they command a more suitable port as both alternatives are too costly. This absence of arbitrage opportunities in the short-run contributes to explain the high autocorrelations in freight rates. Our findings generally confirm the results of Kavussanos (1996a, 1996b, 1997). However, as our analysis is based on daily rather than monthly data, autocorrelation up to the order of 3 is much more pronounced in our study.

Moreover, all α_1 and β_1 coefficients are significant at the 1% level, providing statistical evidence that the GARCH framework is appropriate for modeling the time-varying volatility of freight rates. In contrast to the dry bulk freight market, which strongly depends on the conditional variance of the previous day (high values of the β_1 coefficients), the tanker freight market is comparatively more exposed to current shocks (high values of the α_1 coefficients). The sum of the α_1 and β_1 coefficients is close to one, hence volatility shocks seem to be persistent in the dry bulk freight market.

[9] Searching for the most appropriate model specification, the partial autocorrelation function (results not shown) suggests an order of autoregressive terms to be included of not higher than 3 for all freight rates.

Table VI – Results of Estimated GARCH-X models for Daily Baltic Exchange Indices

This table presents the results of the estimated GARCH-X models for daily Baltic Exchange indices over the period from 1 March 1999 to 18 October 2011. The coefficients denote the following macroeconomic variables: X_1 : slope of the yield curve, X_2 : TED spread, X_3 : world stock market, X_4 : oil, X_5 : wheat, and X_6 : metal. Standard errors are in parentheses. AIC is the Akaike Information Criterion for model selection (Akaike (1974)). LL denotes the value of the log-likelihood function and DW the Durbin-Watson statistic. ARCH LM denotes the p-value for the Lagrange multiplier test for autoregressive conditional heteroscedasticity in the residuals with a time lag of 3 and 10, respectively (Engle (1982)). The coefficient of the GSCI is not reported because the GSCI does not improve the goodness of fit according to the AIC. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Conditional Mean Equation:

$\Delta f r_t = c + \sum_{m=1}^{3} \gamma_m \cdot \Delta f r_{t-m} + \epsilon_t$

Conditional Variance Equation:

$\sigma_t^2 = \omega + \beta_1 \cdot \sigma_{t-1}^2 + \alpha_1 \cdot \epsilon_{t-1}^2 + \sum_{k=1}^{6} \theta_k \cdot X_k(l)$

	BPI		CPI		BCTI		BDTI	
	Normal distribution	t-distribution	Normal distribution	t-distribution	Normal distribution	t-distribution	Normal distribution	t-distribution
Mean								
c	0.001*				0.000	−0.002***	−0.001	
	(0.001)				(0.001)	(0.000)	(0.001)	
γ_1	1.040***	1.057***	0.950***	0.975***	0.514***	0.562***	0.600***	0.610***
	(0.018)	(0.017)	(0.016)	(0.017)	(0.029)	(0.018)	(0.027)	(0.017)
γ_2	−0.174***	−0.203***	−0.085***	−0.161***	0.123***	0.091***	−0.070***	0.037**
	(0.026)	(0.026)	(0.024)	(0.024)	(0.035)	(0.019)	(0.026)	(0.016)
γ_3	−0.088***	−0.075***	−0.089***	−0.055***	0.102***	0.076***		
	(0.020)	(0.017)	(0.017)	(0.017)	(0.029)	(0.015)		

(continued)

Table VI – *Continued*

Conditional Mean Equation: $\Delta fr_t = c + \sum_{m=1}^{3} \gamma_m \cdot \Delta fr_{t-m} + \epsilon_t$

Conditional Variance Equation: $\sigma_t^2 = \omega + \beta_1 \cdot \sigma_{t-1}^2 + \alpha_1 \cdot \epsilon_{t-1}^2 + \sum_{k=1}^{6} \theta_k \cdot X_k(l)$

	BPI		CPI		BCTI		BDTI	
	Normal distribution	t-distribution	Normal distribution	t-distribution	Normal distribution	t-distribution	Normal distribution	t-distribution
Variance								
ω	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	–	–	–	–	–	–	–	–
β_1	0.861***	0.767***	0.835***	0.738***	0.349***	0.433***	0.041	0.367***
	(0.004)	(0.017)	(0.003)	(0.013)	(0.028)	(0.041)	(0.054)	(0.056)
α_1	0.134***	0.226***	0.163***	0.258***	0.088***	0.370***	0.101***	0.429***
	(0.004)	(0.016)	(0.003)	(0.013)	(0.011)	(0.032)	(0.014)	(0.043)
θ_1	0.000***	0.000***	0.000***	0.000***	0.001***	−0.000***		
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)		
θ_2							0.002	0.003
							(0.000)	(0.001)
θ_3					0.002***		0.004***	
					(0.000)		(0.000)	
θ_4	−0.000***						0.003***	
	(0.000)						(0.000)	
θ_5	0.000***							
	(0.000)							
θ_6			0.000***	0.000*	0.001***		0.001***	
			(0.000)	(0.000)	(0.000)		(0.000)	
Diagnostics								
Adj. R^2	0.720	0.720	0.522	0.526	0.286	0.272	0.308	0.305
AIC	−6.655	−6.826	−6.218	−6.440	−6.166	−6.989	−5.345	−5.883
LL	10,967	11,246	10,245	10,612	10,162	11,515	8,812	9,696
DW	1.910	1.940	2.143	2.215	2.316	2.402	2.259	2.283
ARCH LM (3)	0.085	0.707	0.001	0.223	0.992	0.999	0.961	0.977
ARCH LM (10)	0.343	0.827	0.033	0.621	0.999	0.999	0.999	0.999

The AIC recommends that macroeconomic factors should be incorporated into the conditional variance equation but not into the conditional mean equation. Therefore, when modeling the time-variation of freight rate risk, the information content of macroeconomic factors is highest in the conditional variance equation. We further experiment with time lags of the macroeconomic factors both in the conditional mean and in the conditional variance equation, but we do not detect any time-delayed interaction, which could potentially provide further insights on causality. Most important and in strong contrast to the normal distribution, the assumption of a t-distribution implies that the explanatory power of the macroeconomic factors vanishes almost completely, except for the slope of the yield curve and the TED spread. Given that the t-distribution is more appropriate to capture fat tails in freight rate changes, a possible explanation is that the macroeconomic variables are only significant because they exhibit substantial information content in terms of fat tails. In fact, the AIC indicates that the assumption of a t-distribution is considerably better suited to explain conditional volatility than a normal distribution is for all GARCH-X models. This observation is further strengthened by the descriptive statistics in Table II, indicating that it is highly unlikely that freight rates exhibit normally distributed error terms. Finally, if the GARCH specifications are correctly specified, there should be no ARCH effects left in the residuals. Applying Engle's (1982) Lagrange multiplier test, Table VI suggests that only under a normal distribution must the null hypothesis of no ARCH effects up to the order of 3 be rejected for the dry bulk freight market. This result indicates a potential misspecification of the underlying model.

Provided that t-distributed error terms are appropriate to model the time-varying volatility of freight rates, only two macroeconomic factors exhibit significant explanatory power: the slope of the yield curve and the TED spread. Detecting a negative relationship between the term spread and subsequent changes in economic activity with a time lag of 2 to 6 months, several studies have already shown that the slope of the yield curve is a reliable predictor of the future state of the business cycle.[10] However, as we analyze the conditional volatility rather than the level of freight rates, we expect a positive relationship. This notion can be inferred from the theoretical foundations of Stopford's (2009) shipping market supply and demand model: the higher the spread of the yield curve is (indicating good global macroeconomic conditions), the higher the conditional volatility of the underlying freight rates will be. Being significant at the 1% level, the positive coefficient θ_1 confirms this expectation for the conditional volatility of the BPI and the BCI (albeit not for the BCTI). As indicated by the positive coefficient θ_2, the BDTI is more exposed to the TED spread, documenting that the shipping services for crude oil are mainly driven by the higher contemporaneous uncertainty

[10] For example, see Estrella and Hardouvelis (1991), Harvey (1991), Ang et al. (2006) and Aruoba et al. (2006).

of market participants. The higher the TED spread is and thus the higher the market participants' uncertainty about the current state is, the higher will be the resulting freight rate volatility of the BDTI.

6.5.2 EGARCH Model

Stopford's (2009) shipping market supply and demand model constitutes an integral part of explanations for the behavior of freight rates. With respect to freight rates volatility, we expect that large unanticipated shocks will exert a higher impact than small innovations do. Given the convexity of the supply curve, we further anticipate that positive shocks will lead to a higher increase of the conditional freight rate volatility compared to negative shocks of the same magnitude. Nelson's (1991) EGARCH model allows us to test these two hypotheses; the empirical findings are illustrated in Table VII.

Analyzing the conditional mean equation of the EGARCH model, we observe qualitatively very similar results as in the GARCH-X models. The parameter estimations of the conditional mean equation are stable across the different model specifications. The constant of the conditional volatility equation is negative and significant in all cases, suggesting that it is only of marginal importance due to the exponential function. Reinforcing the empirical results of the GARCH-X model, the estimated β_1 coefficients again suggest that the dry bulk freight market is mainly driven by the conditional volatility of the previous day, while the estimated α_1 coefficients indicate that the tanker market is also highly influenced by contemporaneous shocks. Our first hypothesis that larger shocks have a higher influence on the conditional volatility compared to smaller shocks is verified across all EGARCH specifications by the positive δ_1 coefficients (significant at the 1% level). However, with respect to our second hypothesis, we only find mixed results across the market segments. As predicted by Stopford's (2009) shipping market supply and demand model, the results for the tanker market support our hypothesis. Specifically, the significantly positive α_1 coefficients indicate that positive shocks in the tanker market lead to a higher (asymmetric) increase in conditional freight rate volatility than do negative shocks of the same magnitude. In contrast, we cannot confirm this prediction for the dry bulk freight market. More specifically, we do not find an asymmetric effect for the BCI, but even a negative asymmetric effect for the BPI. However, these observations do not necessarily invalidate Stopford's (2009) theoretical model. Instead, another effect in the opposite direction could be at work that cancels out the predicted asymmetric effects of shocks in the shipping industry. An explanation could be deduced from market participants' uncertainty. Specifically, compared to good news, the information content of bad news causes a higher increase of investor uncertainty. The resulting shock outweighs the change that is captured by the supply and demand model which, however, does not incorporate investor uncertainty. As a result, asymmetric effects cannot be detected. In order to further support this notion, Appendix B analyzes the conditional volatility of the MSCI World Stock Market Index using an EGARCH model. Negative innovations exert a higher impact on conditional variance of the stock market than positive innovations do, which strongly supports our explanation. This rationale is also reinforced by the results of Roumpis and Syriopoulos (2009), documenting that the average impact of negative returns of shipping stocks on their return volatility is stronger than that of equivalent positive shocks.

Chen and Wang (2004) and Hui et al. (2008) document similar coefficients, but they do not apply the more appropriate t-distribution. Most important, documenting that negative changes in freight rates have a higher influence on the volatility than positive changes do, the empirical results of Chen and Wang (2004) completely contradict the predictions of Stopford's (2009) shipping market supply and demand model. However, the AIC again indicates that the t-distribution better fits the data across all vessel classes, presumably leading to superior volatility estimates. Furthermore, there is again evidence for potential misspecifications under the assumption of a normal distribution. First of all, in line with the findings by Hui et al. (2008), we also observe ARCH effects in the residuals of dry bulk freight rates. Secondly, although the β_1 coefficients for tanker freight rates seem to be highly significant, their negative values have no meaningful economic interpretation.

Table VII – Results of Estimated EGARCH model for Daily Changes of Baltic Exchange Indices

This table presents the results of the estimated EGARCH models for daily Baltic Exchange indices over the period from from 1 March 1999 to 18 October 2011. Standard errors are in parentheses. AIC is the Akaike Information Criterion for model selection (Akaike (1974)). LL denotes the value of the log-likelihood function and DW the Durbin-Watson statistic. ARCH LM denotes the p-value for the Lagrange multiplier test for autoregressive conditional heteroscedasticity in the residuals with a time lag of 3 and 10, respectively (Engle (1982)). The coefficient of the GSCI is not reported because the GSCI does not improve the goodness of fit according to the AIC. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Conditional Mean Equation:

$\Delta fr_t = c + \sum_{m=1}^{3} \gamma_m \cdot \Delta fr_{t-m} + \epsilon_t$

Conditional Variance Equation:

$\sigma_t^2 = \exp\Big(\omega + \beta_1 \cdot \ln(\sigma_{t-1}^2) + \delta_1 \cdot \Big|\frac{\epsilon_{t-1}}{\sigma_{t-1}}\Big| + \alpha_1 \cdot \frac{\epsilon_{t-1}}{\sigma_{t-1}}\Big)$

	BPI		CPI		BCTI		BDTI	
	Normal distribution	t-distribution	Normal distribution	t-distribution	Normal distribution	t-distribution	Normal distribution	t-distribution
Mean								
c	0.001***		0.001**		−0.001*	−0.001***	−0.002	
	(0.000)		(0.001)		(0.001)	(0.000)	(0.001)	
γ_1	1.016***	1.051***	0.966***	0.983***	0.570***	0.560***	0.601***	0.608***
	(0.015)	(0.017)	(0.016)	(0.018)	(0.023)	(0.016)	(0.020)	(0.016)
γ_2	−0.138***	−0.201***	−0.101***	−0.167***	0.075***	0.091***	0.076***	0.035**
	(0.023)	(0.025)	(0.024)	(0.026)	(0.015)	(0.018)	(0.015)	(0.015)
γ_3	−0.116***	−0.075***	−0.096***	−0.058***	0.099***	0.070***		
	(0.018)	(0.017)	(0.018)	(0.018)	(0.019)	(0.014)		

(continued)

Table VII – *Continued*

Conditional Mean Equation:

$\Delta fr_t = c + \sum_{m=1}^{3} \gamma_m \cdot \Delta fr_{t-m} + \epsilon_t$

Conditional Variance Equation:

$\sigma_t^2 = \exp\left(\omega + \beta_1 \cdot \ln(\sigma_{t-1}^2) + \delta_1 \cdot \left|\frac{\epsilon_{t-1}}{\sigma_{t-1}}\right| + \alpha_1 \cdot \frac{\epsilon_{t-1}}{\sigma_{t-1}}\right)$

Variance								
ω	−0.352***	−0.499***	−0.440***	−0.736***	−13.855***	−3.114***	−10.877***	−2.659***
	(0.013)	(0.053)	(0.012)	(0.056)	(0.641)	(0.336)	(0.986)	(0.352)
β_1	0.979***	0.974***	0.979***	0.961***	−0.546***	0.693***	−0.330***	0.705***
	(0.001)	(0.005)	(0.001)	(0.006)	(0.072)	(0.034)	(0.121)	(0.041)
δ_1	0.222***	0.351***	0.373***	0.619***	0.157***	0.343***	0.149***	0.404***
	(0.006)	(0.024)	(0.008)	(0.046)	(0.014)	(0.028)	(0.012)	(0.043)
α_1	0.002	−0.027*	0.007	0.011	0.017	0.110***	0.012	0.102***
	(0.006)	(0.016)	(0.007)	(0.022)	(0.012)	(0.021)	(0.011)	(0.028)
Diagnostics								
Adj. R^2	0.718	0.720	0.521	0.526	0.270	0.273	0.307	0.305
AIC	−6.620	−6.827	−6.225	−6.457	−6.071	−6.974	−5.269	−5.873
LL	10,908	11,248	10,258	10,639	10,004	11,491	8,685	9,680
DW	1.873	1.930	2.175	2.229	2.426	2.400	2.259	2.280
ARCH LM (3)	0.003	0.826	0.001	0.512	0.987	0.998	0.843	0.992
ARCH LM (10)	0.107	0.978	0.028	0.776	0.999	0.999	0.999	0.999

6.5.3 EGARCH-X Model

So far, both the GARCH-X model and the EGARCH model provide empirical evidence for the theoretical foundations presented in Section 6.2. Searching for the model that is best suited to capture the conditional volatility of both the dry bulk and the tanker freight market, the remaining question of interest is whether both approaches should be considered simultaneously, or whether only one of them exhibits explanatory power in a joint analysis. Specifically, estimating the EGARCH-X model allows us to test for both the presence of asymmetric effects and for the explanatory power of macroeconomic factors in the conditional mean and in the conditional variance equation. Table VIII shows the empirical results.

With respect to the appropriate specification of the conditional mean equation, our findings remain qualitatively similar. Examining the influence of macroeconomic variables as well as the impact of asymmetric effects on the conditional volatility of freight rates simultaneously, we also receive results comparable to the earlier GARCH-X and EGARCH models. For the sake of brevity, we omit a more detailed discussion of the estimated coefficients. We detect remaining ARCH effects for the BPI, the BCI, and the BDTI, indicating potential misspecifications under the normal distribution. Although the β_1 coefficient of the BCTI is significant at the 1% level, the corresponding model seems to be misspecified, given a negative value of 0.120. Again, for all EGARCH-X models the AIC indicates that the t-distribution is more appropriate to capture the time-varying volatility of freight rates than the normal distribution is.

Although Alizadeh and Nomikos (2011) also apply the EGARCH-X model in order to examine the volatility of freight rates, their empirical results are difficult to compare with ours. In contrast to our study, their analysis is based on weekly data, which significantly reduces the number of observations. Moreover, the macroeconomic factor of interest is the term structure of freight rates rather than the slope of the yield curve and the TED spread, respectively. Finally, Alizadeh and Nomikos (2011) do not apply the t-distribution. However, a comparison of the adjusted R-squares indicates that the explanatory power in our model specifications is substantially higher (with values between 27.3% and 72.1%).

As indicated by the high significance levels, our empirical results suggest that both driving forces – macroeconomic variables and asymmetric effects – contribute to explain the conditional volatility of freight rates. In theory, this result should also be valid for the dry bulk freight market. However, depending on the degree of market participants' uncertainty and thus on the underlying observation period, it is possible that there are positive, negative, or no asymmetric effects in the dry bulk freight market.

Table VIII – Results of Estimated EGARCH-X models for Daily Baltic Exchange Indices

This table presents the results of the estimated EGARCH-X models for daily Baltic Exchange indices over the period from 1 March 1999 to 18 October 2011. The coefficients denote the following macroeconomic variables: X_1 : slope of the yield curve, X_2 : TED spread, X_3 : world stock market, X_4 : oil, X_5 : wheat, and X_6 : metal. Standard errors are in parentheses. AIC is the Akaike Information Criterion for model selection (Akaike (1974)). LL denotes the value of the log-likelihood function and DW the Durbin-Watson statistic. ARCH LM denotes the p-value for the Lagrange multiplier test for autoregressive conditional heteroscedasticity in the residuals with a time lag of 3 and 10, respectively (Engle (1982)). The coefficient of the GSCI is not reported because the GSCI does not improve the goodness of fit according to the AIC. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Conditional Mean Equation:

$\Delta fr_t = c + \sum_{m=1}^{3} \gamma_m \cdot \Delta fr_{t-m} + \epsilon_t$

Conditional Variance Equation:

$\sigma_t^2 = \exp\Big(\omega + \beta_1 \cdot \ln(\sigma_{t-1}^2) + \delta_1 \cdot \Big|\frac{\epsilon_{t-1}}{\sigma_{t-1}}\Big| + \alpha_1 \cdot \frac{\epsilon_{t-1}}{\sigma_{t-1}} + \sum_{k=1}^{6} \theta_k \cdot X_k(l)\Big)$

	BPI		CPI		BCTI		BDTI	
	Normal distribution	t-distribution	Normal distribution	t-distribution	Normal distribution	t-distribution	Normal distribution	t-distribution
Mean								
c	0.001***		0.001**		−0.002***	−0.001***	−0.001	
	(0.001)		(0.001)		(0.001)	(0.000)	(0.001)	
γ_1	1.028***	1.053***	0.968***	0.984***	0.555***	0.559***	0.491***	0.611***
	(0.016)	(0.017)	(0.017)	(0.018)	(0.022)	(0.016)	(0.009)	(0.016)
γ_2	−0.151***	−0.204***	−0.101***	−0.169***	0.100***	0.091***	0.125***	0.035**
	(0.024)	(0.025)	(0.026)	(0.026)	(0.017)	(0.018)	(0.016)	(0.015)
γ_3	−0.109***	−0.074***	−0.093***	−0.056***	0.085***	0.070***		
	(0.018)	(0.017)	(0.018)	(0.018)	(0.012)	(0.014)		

(continued)

Table VIII – *Continued*

Conditional Mean Equation: $\Delta fr_t = c + \sum_{m=1}^{3} \gamma_m \cdot \Delta fr_{t-m} + \epsilon_t$

Conditional Variance Equation: $\sigma_t^2 = \omega + \beta_1 \cdot \sigma_{t-1}^2 + \alpha_1 \cdot \epsilon_{t-1}^2 + \sum_{k=1}^{6} \theta_k \cdot X_k(l)$

	BPI		CPI		BCTI		BDTI	
	Normal distribution	t-distribution	Normal distribution	t-distribution	Normal distribution	t-distribution	Normal distribution	t-distribution
Variance								
ω	−0.535***	−0.610***	−0.523***	−0.818***	−10.491***	−3.099***	−0.139***	−3.852***
	(0.027)	(0.070)	(0.019)	(0.068)	(0.234)	(0.326)	(0.008)	(0.513)
β_1	0.965***	0.965***	0.974***	0.955***	−0.120***	0.687***	−0.985***	0.587***
	(0.002)	(0.006)	(0.002)	(0.006)	(0.026)	(0.034)	(0.001)	(0.057)
δ_1	0.240***	0.365***	0.376***	0.631***	0.216***	0.336***	0.009***	0.417***
	(0.008)	(0.026)	(0.008)	(0.047)	(0.013)	(0.028)	(0.003)	(0.045)
α_1	−0.003	−0.027*	0.009	0.013	0.066***	0.110***	0.050***	0.100***
	(0.008)	(0.017)	(0.007)	(0.022)	(0.014)	(0.020)	(0.003)	(0.030)
θ_1	1.919***	1.065**	1.462***	1.301**	8.282***	−3.819***		
	(0.150)	(0.458)	(0.174)	(0.639)	(0.463)	(1.184)		
θ_2							0.430***	11.165***
							(0.033)	(2.134)
θ_3	2.103***		−2.851***		34.172***		4.599***	
	(0.731)		(0.646)		(0.304)		(0.289)	
θ_4					11.817***		−2.775***	
					(0.415)		(0.104)	
θ_5	1.216***							
	(0.380)							
θ_6								
Diagnostics								
Adj. R^2	0.718	0.721	0.521	0.525	0.274	0.273	0.319	0.305
AIC	−6.639	−6.827	−6.233	−6.457	−6.264	−6.977	−5.405	−5.888
LL	10,942	11,250	10,273	10,641	10,325	11,497	8,912	9,706
DW	1.893	1.933	2.175	2.232	2.389	2.398	2.046	2.285
ARCH LM (3)	0.005	0.817	0.002	0.433	0.791	0.999	0.000	0.983
ARCH LM (10)	0.125	0.962	0.039	0.699	0.999	0.999	0.000	0.999

Taken together, the most suitable specification for the heteroscedasticity of freight rates incorporates both macroeconomic variables as well as asymmetric effects simultaneously and applies the t-distribution for the error terms. This main result of our study is of utmost importance for ship owners, operators and bankers alike. One would expect that our proposed specification for the instantaneous volatility of freight rates lays the foundation for better risk management decisions in the maritime industry, e.g. by deriving a more accurate pricing of freight options.

6.6 Conclusion

This study analyzes the volatility structure of freight rates in the dry bulk and tanker freight markets. Justified by the findings of earlier theoretical and empirical studies, we investigate the impact of macroeconomic variables as well as asymmetric effects on the conditional volatility of freight rates by using a GARCH-X model and an EGARCH model, respectively. The question of interest is whether both effects should be considered simultaneously in order to explain the time-varying volatility of freight rates or not. Using an EGARCH-X model, we simultaneously incorporate macroeconomic variables and asymmetric effects one after another and test several combinations in terms of their explanatory power. Moreover, in order to better account for fat tails, we also model t-distributed error terms in addition to using the normal distribution.

Searching for the most suitable specification that is able to capture the time-variation in the volatility of freight rates, three important conclusions can be derived from our analysis. First of all, without any exception, all model specifications indicate that the assumption of a t-distribution is much better suited to explain the conditional volatility than a normal distribution is. Secondly, our analysis suggests that macroeconomic factors should be included in the conditional variance equation, but not in the conditional mean equation. Furthermore, the number of macroeconomic factors that exhibit explanatory power decreases under a t-distribution. We document that the TED spread is highly significant when included in the conditional variance equation of the BDTI, whereas the yield curve seems to have some explanatory power for the volatility of the BPI, BCI, and BCTI. Finally, in contrast to prior studies, we cannot detect asymmetric effects in the dry bulk freight market. However, these effects are strongly pronounced in the tanker freight market. Overall, our empirical findings have important implications for freight rate risk management. Presumably, EGARCH-X models with t-distributed error terms will deliver superior volatility estimates and help to derive a more accurate pricing of freight rate options.

Table A1 – Macroeconomic Variables

Provided by Clarksons Shipping Intelligence Network, this table presents Baltic Exchange freight rate indices, where BPI denotes the Baltic Panamax Index, BCI the Baltic Capesize Index, BDTI the Baltic Dirty Tanker Index, and BCTI the Baltic Clean Tanker Index. The calculations are based on logarithmic differences of the indices. Jarque-Bera denotes the *p*-value of the Bera and Jarque (1980) test for normality. ARCH (3) describes the *p*-value of the Lagrange multiplier test for 3rd order autoregressive conditional heteroscedasticity in the residuals (Engle (1982)).

Statistics	Datastream-code	Details
Global stock market	MSWRLD$	MSCI WORLD U$-Price Index
Oil	LCRINDX	London Brent Crude Oil Index U$/BBL - Price Index
Wheat	WHEATHD	Wheat, No. 2 Hard (Kansas) - cents/bushels
Metal	LMEINDX	London Metal Exchange Index - Price Index
GSCI	GSCITOT	S&P GSCI Commodity Total Return - Return Index
Yield curve	FRTCW10(IR); FRTBW3M(IR)	[US Treasury constant maturities 10 yr - middle rate] -[US Treasury Bill 2nd market 3 months - middle rate]
TED spread	FRTBW3M(IR); BBGBP3M(IO)	[UK interbank 3 months (LDN:BBA) - offered rate] - [US Treasury Bill 2nd market 3 months - middle rate]

Table A2 – Results of Estimated EGARCH Model for the MSCI World Index

Standard errors are in parentheses. *, **, and *** state significance at the 10%, 5%, and 1% level, respectively. AIC is the Akaike Information Criterion for model selection (Akaike (1974)). LL denotes the value of the log-likelihood function and DW the Durbin-Watson statistic. ARCH LM denotes the p-value for the Lagrange multiplier test for autoregressive conditional heteroscedasticity in the residuals with a time lag of 3 and 10, respectively (Engle (1982)).

Conditional Mean Equation:

$$r_t = \gamma_1 \cdot r_{t-1} + \epsilon_t.$$

Conditional Variance Equation:

$$ln(\sigma_t^2) = \omega + \beta_1 ln(\sigma_{t-1}^2) + \delta_1 \left|\frac{\epsilon_{t-1}}{\sigma_{t-1}}\right| + \alpha_1 \frac{\epsilon_{t-1}}{\sigma_{t-1}}$$

	MSCI World Index	
	Normal distribution	*t*-distribution
Mean		
c		
γ_1	0.138***	0.143***
	(0.019)	(0.018)
Variance		
ω	-0.230***	-0.231***
	(0.023)	(0.028)
β_1	0.985***	0.984***
	(0.002)	(0.002)
δ_1	0.110***	0.107***
	(0.013)	(0.016)
α_1	-0.091***	-0.110***
	(0.006)	(0.011)
Diagnostics		
Adj. R^2	0.014	0.114
AIC	-6.648	-6.670
LL	10,958	10,994
DW	2.016	2.026
ARCH LM (3)	0.010	0.016
ARCH LM (10)	0.002	0.049

References

Akaike, H. (1974). "A New Look at the Statistical Model Identification." In: *IEEE Transactions on Automatic Control* 6, pp. 716–723.

Alizadeh, A.H., R. Batchelor, and I. Visvikis (2007). "Forecasting Spot and Forward Prices in the International Freight Market." In: *International Journal of Forecasting* 23, pp. 101–114.

Alizadeh, A.H. and M.G. Kavussanos (2002). "Seasonality Patterns in Tanker Spot Freight Rate Markets." In: *Economic Modelling* 19, pp. 747–782.

Alizadeh, A.H. and N.K. Nomikos (2009). *Shipping Derivatives and Risk Management*. Hampshire: Palgrave Macmillan.

– (2011). "Dynamics of the Term Structure and Volatility of Shipping Freight Rates." In: *Journal of Transportation and Policy* 45, pp. 105–128.

Alizadeh, A.H., N.K. Nomikos, and S. van Dellen (2011). "An Investigation into the Correct Specification for Volatility in the Shipping Freight Rate Markets." Working paper, Cass Business School.

Ang, A., M. Piazzesi, and M. Wei (2006). "What Does the Yield Curve Tell Us about GDP Growth?" In: *Journal of Econometrics* 131, pp. 359–403.

Aruoba, B, F.X. Diebold, and G.D. Rudebusch (2006). "The Macroeconomy and the Yield Curve: A Dynamic Latent Factor Approach." In: *Journal of Econometrics* 131, pp. 309–338.

Bera, A.K. and C.M. Jarque (1980). "Efficient Tests for Normality, Homoscedasticity and Serial Independence of Regression Residuals." In: *Economics Letters* 6, pp. 255–259.

Bessler, W., W. Drobetz, and J. Seidel (2008). "Ship Funds as a New Asset Class: An Empirical Analysis of the Relationship between Spot and Forward Prices in Freight Markets." In: *Journal of Asset Management* 9, pp. 102–120.

Black, F. (1976). "The Pricing of Commodity Contracts." In: *Journal of Financial Economics* 3, pp. 167–169.

Bollerslev, T. (1986). "Generalized Autoregressive Conditional Heteroskedasticity." In: *Journal of Econometrics* 31, pp. 307–327.

Cao, C.Q. and D.B. Nelson (1992). "Inequality Constraints in the Univariate GARCH Model." In: *Journal of Business and Economic Statistics* 10, pp. 229–235.

Chen, Y.-S. and S.-T. Wang (2004). “The Empirical Evidence of the Leverage Effect on Volatility in International Bulk Shipping Market.” In: *Maritime Policy & Management* 31, pp. 109–124.

Ding, Z., R.F. Engle, and C.W.J. Granger (1993). “A Long Memory Property of Stock Returns and a New Model.” In: *Journal of Empirical Finance* 1, pp. 83–106.

Engle, R.F. (1982). “Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation.” In: *Econometrica* 50, pp. 987–1007.

Estrella, A. and G. Hardouvelis (1991). “The Term Structure as a Predictor of Real Economic Activity.” In: *Journal of Finance* 46, pp. 555–576.

Ferson, W.E. and C.R. Harvey (1994). “Sources of Risk and Expected Returns in Global Equity Markets.” In: *Journal of Banking and Finance* 18, pp. 775–803.

Glen, D.R. and B.T. Martin (1998). “Conditional Modelling of Tanker Market Risk Using Route Specific Freight Rates.” In: *Maritime Policy & Management* 25, pp. 117–128.

Glosten, L. R., Jagannathan, R., and D. E. Runkle (1993). “On the Relation between Expected Value and the Volatility of the Nominal Excess Return on Stocks.” In: *Journal of Finance* 48, pp. 1779–1801.

Harvey, C.R. (1991). “The World Price of Covariance Risk.” In: *Journal of Finance* 46, pp. 111–157.

Hui, W., L. Jing, and P.B. Marlow (2008). “An Analysis of Freight Rate Volatility in Dry Bulk Shipping Markets.” In: *Maritime Policy & Management* 35, pp. 237–251.

Kavussanos, M.G. (1996a). “Price Risk Modelling of Different Size Vessels in the Tanker Industry Using Autoregressive Conditional Heteroskedasticity (ARCH) Models.” In: *The Logistics and Transportation Review* 32, pp. 161–176.

– (1996b). “Comparisons of Volatility in the Dry-Cargo Ship Sector.” In: *Journal of Transport Economics and Policy* 30, pp. 67–82.

– (1997). “The Dynamics of Time-Varying Volatilities in Different Size Second-Hand Ship Prices of the Dry-Cargo Sector.” In: *Applied Economics* 29, pp. 433–443.

– (2003). “Time Varying Risks among Segments of the Tanker Freight Markets.” In: *Maritime Economics & Logistics* 5, pp. 227–250.

Koekebakker, S., R. Adland, and S. Sødal (2007). “Pricing Freight Rate Options.” In: *Transportation Research: Part E* 43, pp. 535–548.

Levy, E. (1997). "Exotic Options: The State of the Art." In: ed. by L. Clewlow and C. Strickland. London: International Thomson Business Press. Chap. Asian Options, pp. 65–98.

Nelson, D.B. (1991). "Conditional Heteroskedasticity in Asset Returns: A New Approach." In: *Econometrica* 59, pp. 347–370.

Roumpis, E. and T. Syriopoulos (2009). "Asset Allocation and Value at Risk in Shipping Equity Portfolios." In: *Maritime Policy & Management* 36, pp. 57–78.

Stopford, M. (2009). *Maritime Economics*. New York: Routledge.

Turnbull, S. M. and L. M. Wakeman (1991). "A Quick Algorithm for Pricing European Average Options." In: *Journal of Financial and Quantitative Analysis* 26, pp. 377–389.

Zakoian, J.-M. (1994). "Threshold Heteroskedastic Models." In: *Journal of Economic Dynamics and Control* 18, pp. 931–955.