Essays on Empirical Asset Pricing and Luxury Watches

Dissertation zur Erlangung des Grades eines Doktors der Wirtschaftswissenschaft

eingereicht an der Fakultät für Wirtschaftswissenschaften der Universität Regensburg

vorgelegt von: Siegfried Köstlmeier

Berichterstatter:
Prof. Dr. Klaus Röder
Prof. Dr. Steffen Sebastian

Tag der Disputation:
11.12.2025

Siegfried Köstlmeier

Essays on Empirical Asset Pricing and Luxury Watches

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of
Doktor der Wirtschaftswissenschaft (Dr. rer. pol.)

Advisors:
Prof. Dr. Klaus Röder
Prof. Dr. Steffen Sebastian

University of Regensburg
Regensburg, Germany

Meinen Eltern

Acknowledgements

First of all, I would like to express my sincere gratitude to my supervisor, Prof. Dr. Klaus Röder, for his constant advice, unwavering support, and the opportunities he provided me during my time at the Department of Finance at the University of Regensburg. I am especially thankful for his trust and for giving me the opportunity to pursue my Ph.D. in the first place. I extend my deep gratitude to my second supervisor, Prof. Dr. Steffen Sebastian, for his valuable feedback on my research. The insightful discussions with both of them were essential in shaping the direction of this research and my development as a scholar. I am particularly grateful for the freedom they gave me to pursue my own ideas. Additionally, I am also very grateful to my former colleague Dr. Christian Walkshäusl for many insightful exchanges and discussions.

Moreover, I would like to thank the German Academic Scholarship Foundation (Studienstiftung des deutschen Volkes) for their support and access to their extensive network.

My deepest thanks also go to my parents who shaped who I am today. I am deeply grateful for their unconditional love, care, and unwavering support throughout this journey. They both worked very hard to make it possible that I have been able to pursue and achieve my academic and personal goals. I dedicate this thesis to them.

I extend my gratitude to all my friends, particularly Markus P., Martin L., Patrick M., Simon K., Michael S., Quirin W., Helena W., Petra W., and Josef W., for their constant encouragement, support, and patience. My deep appreciation also goes to Angela G. for her perpetual belief in me. Our journey has been remarkable, but I believe the best is still ahead of us.

November 2025 *Siegfried Köstlmeier*

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Diversification and risk sharing are fundamental principles of modern portfolio theory, asset pricing, and macroeconomics. It is well established since the seminal work of Markowitz (1952) that, under standard asset pricing assumptions, only systematic risk is priced. However, empirical evidence over the past forty decades has uncovered robust return patterns that contradict this classic theory of asset pricing. Prominent and early examples are that value stocks outperform growth stocks (see Basu (1977) and Rosenberg et al. (1985)) or that small stocks outperform big stocks (see Banz (1981)). Nowadays, hundreds of portfolio strategies (often called "factors") to exploit these anomalies have been proposed by academics and practitioners, which is commonly referred to as *"a zoo of new factors"* (see Cochrane (2011) and Harvey and Liu (2016)).

Broadly speaking, the field of asset pricing responded to this development in two ways, as will be explained in more detail below.[1]

The first approach is to continue to adhere to the common assumption in neoclassical finance of rational investors. In accordance with this, investors make decisions with respect to expected utility and apply Bayes' rule to update their beliefs about future outcomes immediately after perceiving new information (which is typically freely available). In this sense, researchers have attempted to explain new empirical findings by extending the discipline's preeminent asset pricing framework, the capital asset pricing model (CAPM) proposed by Sharpe (1964), Lintner (1965), and Mossin (1966). The traditional CAPM framework stipulates that no source of risk other than market exposure should exhibit predictive ability for the cross-section of stock returns. In other words, the market portfolio

[1]See Alquist et al. (2018) for further discussion on this development.

is the central cornerstone of any asset pricing model, and the CAPM states that it is the only one. Empirical evidence in favor of the CAPM is, however, very weak, and the model is rejected, at least since the findings presented in Fama and French (1992) and Fama and French (1993). The proponents of market efficiency as hypothesized by Fama (1970) argue, that the evidence of other variables such as market capitalization (Banz (1981)) or book-to-market ratio (Rosenberg et al. (1985)) to be related to future stock returns simply indicates that the CAPM is misspecified. For that reason, to rationalize empirical findings contradicting the CAPM, other sources of priced risk beyond market risk should be considered in any modern asset pricing framework. Recent developments in models of empirical asset pricing have taken up this challenge, see e.g., Fama and French (2015), Barillas and Shanken (2018), Kelly et al. (2019), Feng et al. (2020), Hou et al. (2021), Kelly et al. (2024), and Cong et al. (2025).

The second approach is to acknowledge the findings from psychological research and to try to incorporate them in asset pricing models.[2] This approach, often referred to as "behavioral finance", emerged from growing evidence that even basic facts about financial markets are hard to reconcile with our traditional, i.e., "rational", models. Shiller (1981) provide evidence that asset prices fluctuate much more than information about their fundamental value. Similarly, De Bondt and Thaler (1985) show that most people overreact to unexpected and dramatic news events. An investment strategy exploiting this behavioral anomaly yields average returns much higher than can be justified by rational measures of risk. Overall, this behavioral approach posits that irrational behavior among investors may have a substantial impact on asset prices and distort them from their fundamental value. Moreover, rational traders who trade on these arising arbitrate opportunities may face risks and costs that limit their ability to remove, or even to correct, this mispricing (see De Long et al. (1990) and Shleifer and Vishny (1997)). Recent developments following this approach of empirical asset pricing have taken up this challenge and proposed "mispricing" related models comprising behavioral factors which are designed to account for investor psychology, see e.g., Stambaugh and Yuan (2017), Asness et al. (2019), and Daniel et al. (2020a).

This divergent development in the field of asset pricing - the "rationalists" on the one side and the "behaviorists" on the other - culminated in the Nobel Memorial Prize in Economic Sciences 2013 to be awarded to three laureates for their very different, partly contradicting, contributions to asset pricing.[3] Since then, however, both sides have been moving closer

[2]See Barberis (2018) for a broad overview of the literature.

[3]Eugene F. Fama was essentially awarded for his evidence that markets are information efficient

together again. Take e.g., Shiller (2017), who emphasizes that the spread and the dynamics of narratives, i.e., stories, particularly those of human interest and emotion, are relevant to economic fluctuations. This resulted in an emerging literature on "narrative asset pricing" in recent years, which infers "narrative news risk factors" by combining textural analysis of newspaper or social media articles with common latent factor analysis, see e.g., Kelly et al. (2021a) and Bybee et al. (2024). Another example is to allow for subjective beliefs in asset pricing models, shifting away from the dominating paradigm of rational agents that know how to form rational forecasts, see e.g., Brunnermeier et al. (2021), Nagel and Xu (2022), and Nagel and Xu (2023).

No matter which side of the two main strands of literature on asset pricing someone is on, risk-based or behavioral models, there is no dispute about the empirical fact that expected stock returns vary over time and are correlated with business cycles (see e.g., Fama and French (1988), Campbell and Shiller (1988a), and Lettau et al. (2008)). Connecting the dots, returns are driven by shocks to expected cash flows and/or shocks to discount rates by definition (see Vuolteenaho (2002)).[4] As emphasized by Cochrane (2011), behavioral theories are thus equivalent with discount rate theories:[5]

> *'It is therefore pointless to argue "rational" versus "behavioral" in the abstract. There is a discount rate and equivalent distorted probability that can rationalize any (arbitrage-free) data. "The market went up, risk aversion must have declined" is as vacuous as "the market went up, sentiment must have increased". Any model only gets its bite by restricting discount rates or distorted expectations, ideally tying them to other data [. . .] And the line between recent "exotic preferences" and "behavioral finance" is so blurred that it describes academic politics better than anything substantive.'*
>
> (John H. Cochrane, 2011)

Some of the apparent split in the field of asset pricing is therefore illusory (see Shiller (2014)). Reconciling the apparently conflicting views is a tough challenge. Nowadays, we have already a much better understanding of how to tackle the problem of testing market efficiency and/or common mispricing, and how cash flow and discount rate shocks carry over from individual stocks to the aggregated market, given our growing set of

in the short-run (see Fama (1970)). Robert Shiller's work refutes this efficient-market hypothesis and focuses investors' irrational behavior. Lars Peter Hansen was awarded for his development on econometric concepts to empirically test asset pricing models. See The Royal Swedish Academy of Sciences (2013) for further information.

[4] In other words, asset prices should equal expected discounted cash flows, as shown in Lucas (1978).

[5] See pp. 1067-1068.

empirical methods and statistical tests. This dissertation also attempts to contribute to this endeavor.

The first three studies of this dissertation contribute to both pillars of the literature on asset pricing. They fill research gaps by investigating stock market anomalies and by analyzing the extent of existing mispricing on the stock market. In a broad sense, the first study accepts the “rationalist” view on anomalies as given and provides novel insights on their interconnectedness through cash flow and discount rate shock spillovers (Chapter 2). Then, the second study takes the “behaviorist” view and accepts that market anomalies may represent *“true”* mispricing, and the contribution of this study is to extend existing models to be able to evaluate the degree of market wide mispricing prevalent in capital markets around the world (Chapter 3). In addition to these more general investigations, the third study turns the focus on a specific anomaly, the momentum effect proposed by Jegadeesh and Titman (1993). Again, both pillars in the field of asset pricing - risk-based models and behavioral models - separately provide reasonable, but sometimes contradictory, arguments for this anomaly. This study contributes to that dispute by presenting a novel approach to decompose momentum returns and to dissect the therein contained risk and mispricing components (Chapter 4).

The last three studies of this dissertation lay the foundation for further research on a nascent and rapidly evolving market: The market for luxury watches. It was not foreseeable at the beginning of this dissertation that luxury watches would become an increasingly serious and attractive investment possibility. Although these three related studies may be seen as only loosely connected to the first three studies of this dissertation, the market for luxury watches is a promising and useful out-of-sample laboratory that could help to shed some light on asset pricing theories in capital markets. This is in line with recent studies that have taken up this challenge by analyzing the sports betting market (Moskowitz (2021)) or cryptocurrencies (Liu et al. (2022)) using standard asset pricing tools and to draw conclusions for asset pricing theories. In this meaning, analyzing the early-stage market for luxury watches may not only help to understand the dynamics of new upcoming markets for other collectibles, but, more important, helps to establish a set of empirical regularities which can be used as stylized facts to asses theoretical asset pricing models. As seen above, this is entirely in line with recent developments in the field of financial research, however, the motivation for financial research studies on luxury watches should be explained in more detail.

Scientific progress and development are hard to predict ex-ante and previously overlooked, new promising areas of research may arise very quickly (see e.g., Kuhn (2012) and Shiffrin et al. (2018)). That was the case in the second half of 2020, when the world was gripped by the COVID-19 pandemic.[6] This period was marked by a historically unique combination of various factors that paved the way for a new field of research on asset pricing, which will be discussed in chapters 5 to 7 of this dissertation. In summary, these incidents were essentially the interplay between historically low interest rates, low inflation rates, a surge in household savings accompanied by a decrease in consumption spending, and governmental enacted fiscal stimulus around the world which transferred billions of dollars to taxpayers. This paved the way for investors to start looking for alternative sources of yield beyond stocks, bonds, real estate, private equity/debt, and hedge funds. They have found these promising investment opportunities in the area of collectibles, and luxury watches are among the most sought-after collectibles in recent years (Boston Consulting Group (2023)).[7]

In more detail, interest rates on U.S. government securities at 10-Year maturity declined since their peak at 15.32% in the early 1980s to a minimum of 0.62% in July 2020. Similarly, average long-term government bond yields for 19 countries in the euro area decreased from 15.44% in September 1981 to a level of zero until the late 2010s and were even negative at the end of 2020. The decline in interest rates over decades has occurred in many countries around the world, and inflation rates have also plummeted alongside this development. Global inflation rates plunged from an average 7.99% p.a. in the 1980s to a meager 2.54% p.a. in the 2010s.[8] Strict lockdown policies resulted in drastic and sustained reductions in air traffic, and people could no longer spend their money on traveling and other leisure activities even local to the place of residence (see Ludvigson et al. (2021)). The U.S. personal saving rate surged to an unprecedented 31.8% in April 2020, and within the euro area, increased to 20.4% in Q1 2021. With the restrictions largely lifted by 2022, the saving rate in the euro are returned to its pre-pandemic average (around 13.0% since 1999), but has, however, increased again over the last two years.[9] Eichenbaum et al. (2021) show that absent containment measures, average consumption falls by about 7% in the first year of an epidemic, and with optimal containment, average consumption even falls by 22%. In addition, the U.S. government enacted fiscal stimulus during that time distributed approximately $814 billion to taxpayers and a significant

[6] For an overview of its impact on the worldwide economy, see Brodeur et al. (2021).

[7] Other popular alternatives in recent years are cryptocurrencies and non-fungible tokens, see Cong et al. (2021) and Corbet et al. (2023).

[8] All data is retrieved from FRED.

[9] Data for the euro area is retrieved by Eurostat (QSA).

portion of that was used for investments (see Greenwood et al. (2023)).

Especially the rise of peer-to-peer online platforms for luxury watches like Chrono24.com (`https://www.chrono24.com`) created ongoing public awareness for the watch market and helped to disseminate that individuals are able to buy and sell luxury watches as convenient as never before. Founded 2003 in Germany, approximately 3,500 commercial dealers (mostly jeweler's) from over 100 countries offer regularly more than 500,000 luxury watches, comprising an aggregated value of roughly 6 billion Euro (see Chrono24 (2025)).[10] Other platforms like WatchCharts Analytics additionally provide the much-needed transparency for luxury watch investors. They collect and analyze millions of data points to determine market prices for over 100 brands and 25,000 watches (see Analytics (2025)). This further attracts new collectors and investors alike and their key benefit is to improve liquidity, price discovery, and transaction performance for luxury watches. In addition, these platforms enable us to gather price data on luxury watches and to analyze their risk- and return-properties, and the papers presented in this dissertation are the first to do so.

Two important questions arise: Why is this relevant to the field of asset pricing and why should the contributions presented in this dissertation be important for the literature? There are three reasons for this.

First, based on the insights of Edmans (2025) from 1,000 manuscript rejections as a managing editor for the *The Review of Finance*, presenting even simple summary statistics for a brand-new data set can be interesting. Given an interesting data set, however, one must *"ask interesting questions with it"* (p. 425). In this sense, Chapters 5 to 7 are the first studies to analyze risk and return properties of luxury watches. Far beyond simple, descriptive statistics, we analyze the diversification potential of luxury watches and address the key question if any benefits stem from an increase in portfolio returns, a reduction of risk, or both together. Because the market for luxury watches provides a useful out-of-sample laboratory that could help to shed some light on asset pricing theories in capital markets, we then analyze this market through the lens of asset pricing. We construct watch counterparts for well-known price- and market-related return predictors in the stock market and document the existence of multiple anomalies in the luxury watch market.

Second, a review of the literature shows that other collectibles are broadly studied and have received considerable attention in finance research. This includes art (Goetzmann

[10] Although private sellers are permitted, more than 90% of all sellers are professional dealers.

(1993), Campbell (2008), Korteweg et al. (2016)), diamonds (Ariovich (1985), Renneboog and Spaenjers (2012) and Auer and Schuhmacher (2013)), fine wines (Krasker (1979), Dimson et al. (2015), see Le Fur and Outreville (2019) for a review of the literature) and cars (Martin (2016), Laurs and Renneboog (2019), Le Fur (2023)), as well as less common collectibles such as antique furniture (Rush (1968)), coins (Kane (1984)), timber (Redmond and Cubbage (1988)), antique firearms (Avery and Colonna (1987)), Stradivarius violins (Ross and Zondervan (1989)), photographs (Pompe (1996)), stuffed animals (Burton and Jacobsen (1999)), sculptures (Locatelli-Biey and Zanola (2002)), postage stamps (Dimson and Spaenjers (2011)), whisky (Moroz and Pecchioli (2019)), non-fungible tokens (Dowling (2022)), LEGO sets (Dobrynskaya and Kishilova (2022)), Magic the Gathering game cards (Langelett and Wang (2023)), comics (Bocart et al. (2023)), and most recently, video game skins (Reichenbach (2025)). This continuous series of studies clearly emphasizes the ongoing interest of financial research in collectibles, yet very little is known about luxury watches. This dissertation fills this gap in the literature.

Third, the findings presented in Chapters 5 to 7 have implications for investors, wealth managers, and luxury watch dealers. Given our research findings, they are provided with a precise understanding of the risk and returns of the various luxury watch series in which they could invest. The latest available Deloitte Art & Finance Report 2023 states that the wealth of ultra-high-net-worth individuals associated with art and collectibles was already an astonishing \$2.174 trillion in 2022 and is expected to be \$2.861 trillion in 2026, which highlights that an increasing number of people are willing to invest in these alternative investment classes (see Deloitte (2023)). Our studies on luxury watches are therefore helpful for related advisors to tilt portfolios of luxury watch investors towards specific investment choices as proxied by our extensive set of analyzed luxury watch characteristics.

The remainder of this doctoral thesis is structured as follows. In the rest of the Introduction, the research papers are briefly summarized with respect to motivation, research questions, data and method, and major results.

Table 1.1:
Overview of the publications with the corresponding chapter, title, and current publication status.

Chapter	Title	Publication	
		Status	Journal (VHB Media Rating 2024)
2	What Drives Anomaly Returns? European Insights and Spillover Effects of Cash Flow and Discount Rate Shocks	Under review	The European Journal of Finance (B)
3	Pricing and Mispricing of Accounting Fundamentals: Global Evidence	Published	The Quarterly Review of Economics and Finance (B)
4	The Relevance of Risk, Mispricing, and Optionality in Momentum Returns	Under review	The Quarterly Review of Economics and Finance (B)
5	Diversification Benefits of Luxury Watches and Day-of-the-week Effects in a Seven-Day Traded Market	Under review	Journal of Banking & Finance (A)
6	The Global Market for Luxury Watches and Asset Pricing	Under review	The Review of Financial Studies (A+)
7	Look at my Watch! Continuous Information and the Momentum Effect in the Market for Luxury Watches	Under review	Financial Markets and Portfolio Management (B)

Table 1.1 provides an overview of these papers, the current publication status, and the assignment to the following chapters. Chapters 2 to 7 constitute the core of this doctoral thesis and present six independent research papers. Chapter 8 concludes.

What Drives Anomaly Returns? European Insights and Spillover Effects of Cash Flow and Discount Rate Shocks

This article contributes to the strand of literature on analyzing cash flow and discount rate shocks among stock returns. Nowadays, hundreds of factors are proposed in the literature to explain the cross-section of expected stock returns (see e.g., Harvey and Liu (2016)). Their economic importance seems to be huge, as documented by Lochstoer and Tetlock (2020), who show that non-market factors account for 85% of the variance in the stochastic discount factor as implied by the model proposed in Fama and French (2015).

For this reason, it is not only important to analyze cash flow and discount rate components among the returns at the firm-level, which is the main interest in earlier studies, e.g., in Campbell (1991) or Vuolteenaho (2002). Rather, analyzing the cash flow and discount rate components among anomaly portfolios themselves seems to be a more promising endeavor. Lochstoer and Tetlock (2020) is the first to provide evidence for the dominance of cash flow shocks among U.S. anomaly portfolios. As with any findings in empirical research, this could by the result of data snooping in the sense of Lo and MacKinlay (1990) and therefore be sample-specific. We address this concern by independently examining the decomposition of anomaly returns in a large sample of international equity markets. In doing this, this study is, however, more than a replication of previous findings using a different data set. While Lochstoer and Tetlock (2020) laid the methodological foundation on how to decompose anomaly portfolio returns in the sense of Campbell (1991), the study mainly focuses on static correlation metrics of cash flow and discount rate shocks with shocks to the market portfolio or shocks to macroeconomic variables. While this is important, the behavior of shocks could be time-varying. We fill this gap in the literature and extend these previous findings by applying a time-varying vector auto-regression (TVP–VAR) based extended joint connectedness framework as proposed in Balcilar et al. (2021) to estimate spillover effect of these cash flow and discount rate shocks.

Our main findings are as follows: In line with Lochstoer and Tetlock (2020) for the U.S., cash flow shocks tend to dominate the returns of portfolios representing the size, value, investment, profitability, and momentum anomaly. Their contribution to the return variance of these anomaly portfolios is within the range from 66.64% for the value anomaly, and 82.65% for the size anomaly. In contrast, discount rate shocks contribute less than 16.54% (momentum) among all anomalies. On average, cash flow shocks also explain 89.71% of market-adjusted return variance at the firm-level, while discount rate shocks are a less important component. Interestingly, while some of the cash flow and discount rate shocks are correlated with shocks to macroeconomic variables, correlations are at best modest in magnitudes, ranging from only -0.22 to 0.15. In summary, we find that anomaly portfolios tend to have a common component located in their cash flow news, but this unobserved component seems to be uncorrelated with shocks in standard macroeconomic variables. Applying a time-varying vector auto-regression (TVP–VAR) based extended joint connectedness framework proposed in Balcilar et al. (2021) for these cash flow and discount rate shocks, we find that, on average, 40.22% of a shock in one of anomaly cash flow or discount rate news spills over to all others. The largest shocks are obtained within discount rates, whereas their spillovers to the cash flow shock category seems to be modest. Despite the dominance of cash flow shocks among anomaly portfolio returns, discount rate shocks seem to play a more important role in carrying shocks from one anomaly to the others. These findings, overall, provide further guidance for the development of asset pricing theories, as shown in Lochstoer and Tetlock (2020).

Pricing and Mispricing of Accounting Fundamentals: Global Evidence

This article of the dissertation contributes to the strand of literature analyzing stock market mispricing. We build upon the insights of Nichols et al. (2017) who document a strong relation between the absolute share price and accounting fundamentals. At least 63% of the cross-sectional share price variation of U.S. firms can be traced back to variations in the book value of equity, income, dividends, and growth in operating income. The valuation model therefore directly addresses the concerns of Cochrane (2011) that the field of asset *pricing* has in fact evolved into asset *expected returns.*

There are two important caveats to be considered. In order to overcome these limitations, this paper is not just a replication of previous studies using an international, non-U.S. data set, but also provides a theoretical extension of the model presented in Nichols et al. (2017).

First, the original model only considers firms whose fiscal year-end is December 31. This would discard a large portion of firm observations around the world. As a gauge of magnitude, 70% of Japanese firms prefer the end of March as fiscal year-end and only 38% of Asia-Pacific firms chose December 31. Their inclusion, however, is not trivial because quarterly accounting data is generally not available for non-U.S. firms prior to 2001.

Second and more important, what the original model denotes as "dividends", is actually an indirectly measured residual figure based on clean-surplus accounting assumptions. In the manner of Ohlson (2005), "dividends" is unbiased only if the expected equity transactions are neutral or irrelevant to prospective new shareholders. This is, however, not the case for firms conducting share repurchases, which have generally become more prevalent in recent years. We observe an aggregated buyback volume of €833.08 billion in Europe (excl. United Kingdom) in the thirty-one years from 1990 to 2021, similar to Fried and Wang (2021). Of that volume, €475.89 billion were repurchased within the most recent decade, and the original valuation model of Nichols et al. (2017) would ignore this form of wealth transform to shareholders. In consequence, we adjust the model to explicitly controlling for share repurchases and to be able to include all firms in our global sample, no matter their respective fiscal year-end.

We investigate on average 4,542 firms among 21 global, developed non-U.S. equity markets from June 1990 to June 2021. Our extended version of the model is able to explain 81% of global firms' share price variation. More important, the parsimonious cross-sectional fundamentals-based valuation model links share prices to publicly available accounting fundamentals, so any deviations of observed share prices from the model's derived share prices is helpful to identify mispricing. We find that a related portfolio strategy generates a highly significant 0.56% p.m. and our extensive analysis indicates that this return actually exploits mispricing opportunities rather than being a reward for facing risk exposure. This conclusion is the result of four separate empirical findings.

First, we document a large post-publication return decline amounting to 0.43% p.m., which is expected in case of (risk-less) arbitrage opportunities. This return drop is entirely rooted in the easily exploitable long portfolio leg comprising undervalued firms. Second, risk-based explanations are less likely based on fast portfolio transitions over consecutive portfolio sorts, implying a rather implausible quick change in the underlying source of risk. Third, our portfolio strategy is most profitable in times of high investor sentiment (0.41%) and generates only 0.17% in times of low sentiment. Last, we observe that our strategy return is even higher (0.71%) for stocks that represents risk for arbitrageurs and thus prevents them from correcting the mispricing.

Overall, these findings are consistent with the notion that fundamental information is only gradually incorporated into share prices by investors. Besides the ability of our extended valuations model to capture firm-level mispricing, it also correctly identifies overall, market-wide times of mispricing because of weaker mapping of firm fundamentals into stock prices.

The Relevance of Risk, Mispricing, and Optionality in Momentum Returns

This study extends our previous insights on stock market mispricing and sets an example based on the momentum effect. In their seminal paper, Jegadeesh and Titman (1993) demonstrate that buying past winner stocks and selling past loser stocks yields positive returns. The robust and large returns generated by momentum strategies are documented among a vast number of asset classes, e.g., mutual funds (Carhart (1997)), commodity futures (Miffre and Rallis (2007)), corporate bonds (Jostova et al. (2013)), cryptocurrencies (Liu et al. (2022)), as well as government bonds and currencies (Asness et al. (2013)). The effect seems to present a striking contradiction to the weak form of market efficiency hypothesized by Fama (1970) and thus *"momentum is a hard sell for a world of rational* *pricing [...]"* (Fama and French (2020), p. 1894). However, the origin of the momentum anomaly is still debated.

Unsurprisingly, in consideration of the two main pillars in the field of asset pricing described above, two main strands of literature - risk-based models and behavioral models - provide reasonable arguments for the momentum effect (see Wiest (2023) for an overview of the literature). Rational and behavioral asset pricing models, however, imply a tug of war in expected returns for speculative stocks because related characteristics (e.g., small, young, lottery-like, or close to distress) overlap with stocks perceived as being risky (Baker and Wurgler (2007); Birru (2018)). These conflicting predictions of rational and behavioral models may confound previous empirical tests examining the momentum effect using standard asset-pricing models. Our paper contributes to bridge this gap and mitigate the academic dispute as our findings imply that momentum returns are essentially driven by both, mispricing *and* risk.

We build upon the insight that momentum strategies are punctuated with occasional crashes (Daniel and Moskowitz (2016)) and develop a novel approach to decompose momentum returns into a risk-, mispricing-, and option-component.[11] Again, this paper is not just an empirical repetition using established econometric methods. We unify the findings proposed in Daniel and Moskowitz (2016) with the approach proposed in Birru (2018) and present a novel return-decomposition approach.

To begin with, we demonstrate the validity and robustness of our decomposition approach in multiple ways. We conduct simulations to assess statistical properties and the power of the approach by means of bootstrap simulations. To evaluate the econometric robustness of our regression-based approach, we also conduct quantile and ridge regressions (see Koenker and Bassett (1978) and Koenker (2005)), as well as a Bayesian framework based on the Markov Chain Monte Carlo method of Chib (1998) to account for possible structural changes in the distribution of momentum returns (see Geman and Geman (1984), Gelfand and Smith (1990), Kass and Raftery (1995), Chib (1995), and Chib (1998)).

After having established the suitability of our novel econometric approach, we decompose the returns of 28 equity momentum strategies into a risk-, mispricing-, and option-component in our empirical analysis. We consider the U.S. market for the period July

[11] Daniel and Moskowitz (2016) emphasize that momentum crashes are best described by on option-like behavior. In times of a crash, the momentum strategy behaves as if it is effectively short a call option on the market portfolio. As a gauge of magnitude, the momentum strategy lost 45.21% in April 2009 during the global financial crisis.

1972 to December 2016 and find that the average return of 0.66% p.m. generated by the standard up-minus-down (UMD) momentum strategy in the sense of Jegadeesh and Titman (1993) consists of a 0.29% risk-component (t-statistic: 1.94), a 0.50% mispricing-component (t-statistic: 6.35), and a -0.13% option-component (t-statistic: -3.57). Looking at all 28 different variants of momentum strategies, we find that the risk-component accounts on average for 56.66% of momentum returns, followed by a 29.63% mispricing-component, and a remaining 13.71% option-component.

In conclusion, our findings are important for the overall understanding of the momentum effect and offer explanations for contrary results in previous studies. Take e.g., the contrary results for the relation between the default spread between U.S. corporate bonds and momentum returns. Chordia and Shivakumar (2002) report a strong negative relation while Griffin et al. (2003) fail to confirm this evidence. We detect that the relation is almost entirely driven by the risk-component of the short momentum portfolio leg. Separating risk- and mispricing in momentum seems to be crucial for our understanding of this phenomenon, especially since decades of intensive research did not result in a commonly accepted explanation. The relation between momentum and market illiquidity as documented in Avramov et al. (2016) is fully subsumed by the mispricing-component of momentum. This finding supports the view in Huber (2022) that strategies to exploit mispricing are difficult to implement in illiquid markets. The negative relation between momentum and aggregated market volatility is fully reflected in its risk-component. Momentum returns are, at least partially, an actual risk premium for market volatility risk, which explains the superior performance of volatility-scaling momentum strategies proposed in Barroso and Santa-Clara (2015), Daniel and Moskowitz (2016), and Wang and Yan (2021).

As usual for a novel econometric contribution, this paper may leave the reader with more questions than answers. After controlling for the mispricing-component, the "risk" in momentum returns seems to be an even greater puzzle than previously thought. Much work is left to identify the sources of risk reflected in the risk-component and how these risks are transmitted into momentum returns.

Diversification Benefits of Luxury Watches and Day-of-the-week Effects in a Seven-Day Traded Market

This paper is the first to analyze the diversification potential of luxury watches and the key question if benefits stem from an increase in portfolio returns, a reduction of risk, or both together.

Why is this important and relevant? The latest available Deloitte Art & Finance Report 2023 reveals that the investments of ultra-high-net-worth individuals only associated with art and collectibles amount to $2.174 trillion. Approximately 63% of surveyed wealth managers have already integrated collectibles into their wealth management offering, and 10.9% of client allocation is associated with related investments.[12] From a broader perspective, real personal consumption expenditures for jewelry and watches in the USA exceed $100 billion (measured in 2024-$, from FRED) in every year since 2021, and the dramatically increase in both desirability and global demand for Rolex watches in recent years is now even discussed in a Harvard Business School case study (Chung (2021)).[13] Given these numbers, it is incomprehensible that luxury watches as a financial investment have been entirely overlooked by academics. We fill this huge gap in the literature and examine important questions that arise to any investor when being confronted with new investment prospects: Do luxury watches provide additional diversification benefits beyond stocks, bonds, and gold, and if so, are there potential day-of-the-week effects that should be accounted for when buying or selling luxury watches?

Our novel data set comprises daily price data for luxury watches from WatchCharts Analytics from 01/01/2017 to 09/30/2024. According to Morgan Stanley's annual watch report (see Müller (2024)), the top five leading Swiss manufactures in terms of worldwide retail market share 2023 are Rolex (30.3%, including their brand Tudor), Cartier (7.5%), Omega (7.5%), Patek Philippe (5.6%), and Audemars Piguet (4.9%). Reflecting their economic relevance, our sample comprises related indices for these six brands covering a combined market share of 55.8%.

First, we find that some of our analyzed luxury watch indices generate quite large returns. The average annualized return of Rolex (6.94%), Patek Philippe (10.61%), and Audemars Piguet (10.81%) is close to the performance of U.S. stocks (9.28%). We observe that return volatility of luxury watches is remarkably low and just about one fifth of stock market volatility, thus quite the same as Treasury bills. Taken together, the annualized Sharpe ratios of Audemars Piguet (1.55), Patek Philippe (1.29), and Rolex (1.26) vastly exceeds the Sharpe ratio of the U.S. stock market (0.71) which is somehow surprising given its remarkable well performance in recent years. Unsurprisingly, our mean-variance

[12]See Deloitte (2023). An ultra-high-net-worth individual is someone with a net worth of at least $30 million.

[13]Beyond wealthy households, several studies indicate that between a quarter and a third of all adults in Western countries define themselves as collectors (see Belk (1988) and Pearce (1995)).

spanning tests as proposed in Huberman and Kandel (1987), Ferson and Keim (1993), and Kan and Zhou (2012) demonstrate that investments in Audemars Piguet, Patek Philippe, and Rolex watches shift the tangential portfolio comprising stocks, bonds, and gold towards higher risk-adjusted returns. With no short sales allowed, the Sharpe ratio doubles to 1.23 when adding luxury watches to the benchmark portfolio of stocks, bonds, and gold, because return volatility almost halves to only 2.35% p.a. Estimated alphas are a remarkable 5.28% p.a. in case of Rolex and exceed even 10% p.a. in case of Patek Philippe and Audemars Piguet.

Second, given that our results clearly show that investors benefit from a diversification potential of luxury watches, we address the question that arises when implementing the actual watch investment: When is the best time, in the sense of which day of the week, to buy them? Highest returns are observed on Wednesdays with 2.42 bps, which is mostly driven by the economically large and statistically highly significant Wednesday return of 3.51 bps for Rolex. In general, returns tend to be higher around the mid of a week with average returns of 2.12 bps on Tuesdays and 1.55 bps on Thursdays (both driven by high returns of Audemars Piguet and Rolex). On Saturdays and Sundays, returns for the six brand indices are on average an insignificant and small 1.12 bps, resp. 0.18 bps. We further apply an ARMA-EGARCH(-M) model (see Engle (1982), Nelson (1991)) to analyze these daily differences in the returns and volatilities of luxury watch portfolios as a robustness test which confirms that returns on Sundays are lower than on other days of the week.

What explains lower returns and volatility of most luxury watch brand portfolios on Sundays? We address this question by (1) looking at patterns in the daily global sales volume on eBay for Rolex models, (2) evaluating retail investor attention proxied by Google Trends web search volume for the term “Rolex”, and (3) analyzing the timestamps of 2,000 watch listings on eBay. Using k-means to identify clusters for the specific time of the day an offer is published or updated, we notice a clear pattern: Most listings are edited on weekdays at 09:03 a.m. and 05.03 p.m., which basically matches the beginning and end of a typical business day. Overall, returns for luxury watches are lower on Sundays because professional dealers typically do not update offers on that day, which is typically a day of rest in most western countries.

The Global Market for Luxury Watches and Asset Pricing

In Chapter 6 we analyze a broad sample of 27,289 hand collected watch-month observation from the world's largest peer-to-peer marketplace for luxury watches Chrono24.com between June 2010 and March 2022 through the lens of asset pricing. This unique and novel data set opens new possibilities to test theories of cross-sectional asset pricing anomalies, so we are the first to test 30 characteristics related with the categories size, value, momentum, and volatility in the cross-section of 345 distinct luxury watches from 20 brands.

Why is this endeavor important and relevant to the field of finance? First, the market for luxury watches provides a useful out-of-sample laboratory that could help to shed some light on asset pricing theories in capital markets. Recent studies have taken up this challenge by analyzing the sports betting market (Moskowitz (2021)) or cryptocurrencies (Liu et al. (2022)) using standard asset pricing tools. Second, recent findings emphasize that the economic influence of wealthy households (i.e., the typical buyers of luxury items) seems to be large enough to shape the markets. Bali et al. (2023) show that the highly skewed distribution of household wealth explains the anomalous negative risk-return relation among high-volatility stocks,[14] providing evidence that few individual investors affect equilibrium asset prices and can contribute to asset pricing puzzles.[15] To our surprise, luxury watches as investment have been completely ignored in top-tier finance and economics journals so far. In this paper, we fill this gap in the literature.

We find that the characteristics size, reversal, short-term momentum, and MAX generate significant difference returns among zero-investment quintile portfolio strategies. Both the k-FWER test method by Lehmann and Romano (2005) and an F-test for the joint significance provide evidence that our results are unlikely to generate by chance. In accordance with the findings of Stambaugh et al. (2015) on the asymmetric pricing effect of sentiment, we find that sentiment-related variation in their performance is mainly due to their short positions. Overall, our results are in favor of a mispricing related interpretation and that the strategies reflect a mispricing commonality across luxury watches.

While this study reveals new evidence consistent with mispricing as at least a partial explanation for our studied watch-counterparts to prominent equity anomalies, we do

[14]Also known as the idiosyncratic volatility puzzle, see Ang et al. (2006).

[15]Their capability to significantly influence the market is accompanied by the fact that retail investors in general have become increasingly important players in financial markets, see Eaton et al. (2022).

not aim to find complete explanations for each of them. Instead, studying the early-stage market for luxury watches helps us to establish a set of empirical regularities which can be used as stylized facts to asses theoretical asset pricing models, and, even more important, helps us to understand the dynamics of new upcoming markets for other collectibles, too.

Look at my Watch! Continuous Information and the Momentum Effect in the Market for Luxury Watches

The momentum effect is one of the most robust and central asset pricing anomalies. In Chapter 4 of this dissertation, we developed a novel decomposition approach to dissect equity momentum returns into a risk-, mispricing-, and option-component, and we construct the luxury watch counterpart of the momentum strategy in Chapter 6. Given the broad focus of Chapter 6 on analyzing an extensive set of thirty characteristics to predict the cross-section of luxury watch returns, one central question has not yet been addressed: Why does the momentum effect occur in the market of luxury watches at all? We fill this gap in the literature with this study and analyze the underlying economic mechanisms of the luxury watch market momentum effect.

We consider 124 luxury watch indices of 26 brands from WatchCharts.com and their daily returns for the period 06/30/2017 to 09/30/2024. Similar with other asset classes, we document a strong momentum effect generating a highly significant return of 1.25% per month.

We find that the inattentiveness of investors to continuously arriving information during the momentum formation period drives momentum returns. This is known as the frog-in-the-pan hypothesis which originates from limited investor attention (see Da et al. (2014)). According to the frog-in-the-pan anecdote, a frog will jump out of a pan containing boiling water since the dramatic temperature change induces an immediate reaction. Conversely, if the water in the pan is slowly raised to a boil, the frog will underreact and perish.

Using bivariate independent portfolio sorts, we find that momentum returns decrease from a highly significant 1.67% for luxury watches with continuous information during their formation period to an insignificant -0.38% for watches with discrete information, but similar cumulative formation-period returns. Our robustness tests further differentiate between information discreteness (Da et al. (2014)), which is motivated by limited attention, and return consistency (Grinblatt and Moskowitz (2004)), whose motivation lies with the disposition effect. Overall, our battery of empirical tests indicates that indeed limited attention is the core economic channel for the return predictability of continuous information. In conclusion, our results suggests that this mispricing related channel for momentum already documented among stocks (see Da et al. (2014)) is also prevalent in driving momentum strategy returns in the market for luxury watches.

Chapter 2

What Drives Anomaly Returns? European Insights and Spillover Effects of Cash Flow and Discount Rate Shocks

This research is the result of a single-author project. The paper is currently under review in *The European Journal of Finance*. The journal ranking is *B* according to the VHB Publication Media Rating 2024.

Abstract

We decompose the returns of the size, value, investment, profitability, and momentum anomaly into cash flow and discount rate shocks. Cash flow shocks explain between 66.64% and 82.65% of these anomaly portfolio returns for firms located in the European Monetary Union, and explain 89.71% of market-adjusted return variance at the firm-level, while discount rate shocks are a less important component. We show that (i) cash flow shocks show a common component among anomaly portfolios, (ii) cash flow shocks show little relation to the business cycle, and (iii), discount rates are the main channel for transmitting spillover effects to other anomalies.

Keywords: Asset Pricing · Return Decomposition · Cash Flow Shocks · Discount Rate Shocks

JEL classification: G12 · G14 · G15.

2.1 Introduction

The finance profession has been on a multiple decade-long quest to identify factors that explain the cross-section of expected stock returns. Hundreds of such factors are nowadays proposed in the literature (see e.g., Harvey and Liu (2016)). Their economic importance seems to be huge, as documented by Lochstoer and Tetlock (2020), who show that non-market factors account for 85% of the variance in the stochastic discount factor as implied by the model proposed in Fama and French (2015).

To gain insights into the economic sources of these non-market factors, Lochstoer and Tetlock (2020) propose a method for decomposing anomaly portfolio returns into cash flow (CF) and discount rate (DR) shocks as in Campbell (1991). Their findings imply that CF news explain most of the variation in anomaly returns, and, more interestingly, CF news continue to dominate in explaining the returns of the mean-variance efficient (MVE) combination of anomaly portfolios.

The motivation for this research study can easily be visualized using the European non-market factors SMB (size), HML (value), RMW (profitability), and CMA (investment), which are obtained from Kenneth French's data library.

Figure 2.1 shows the 3-year rolling correlation for all pairwise factor returns from July 1991 to June 2025. While the unconditional average among all pairwise factor correlations is close to zero (-0.09), there are notable exceptions, e.g., the robust negative average correlation of -0.66 between HML and RMW, or the positive average correlation of 0.47 between HML and CMA. Recent studies provide further evidence that the cross-section of anomaly portfolio returns is indeed auto-correlated (see Ehsani and Linnainmaa (2022)). More interesting, the time-series variation of 3-year rolling correlation is huge. For instance, the correlation between size and value returns varies between -0.35 and 0.48 until June 2025. What drives these anomaly returns, and is there a common component to cause their returns to be correlated? What causes the time-series variation in their covariance structure? This study provides answers and fills these gaps in the literature.

In this study, we contribute to the literature by studying the importance of CF and DR shocks in anomaly portfolios outside the United States. As with any findings in empirical research, the dominance of cash flow shocks among anomaly portfolios could by the result of data snooping in the sense of Lo and MacKinlay (1990) and therefore be sample-specific. We address this concern by independently examining the decomposition of stock returns in a large sample of international equity markets. Our sample comprises the ten countries

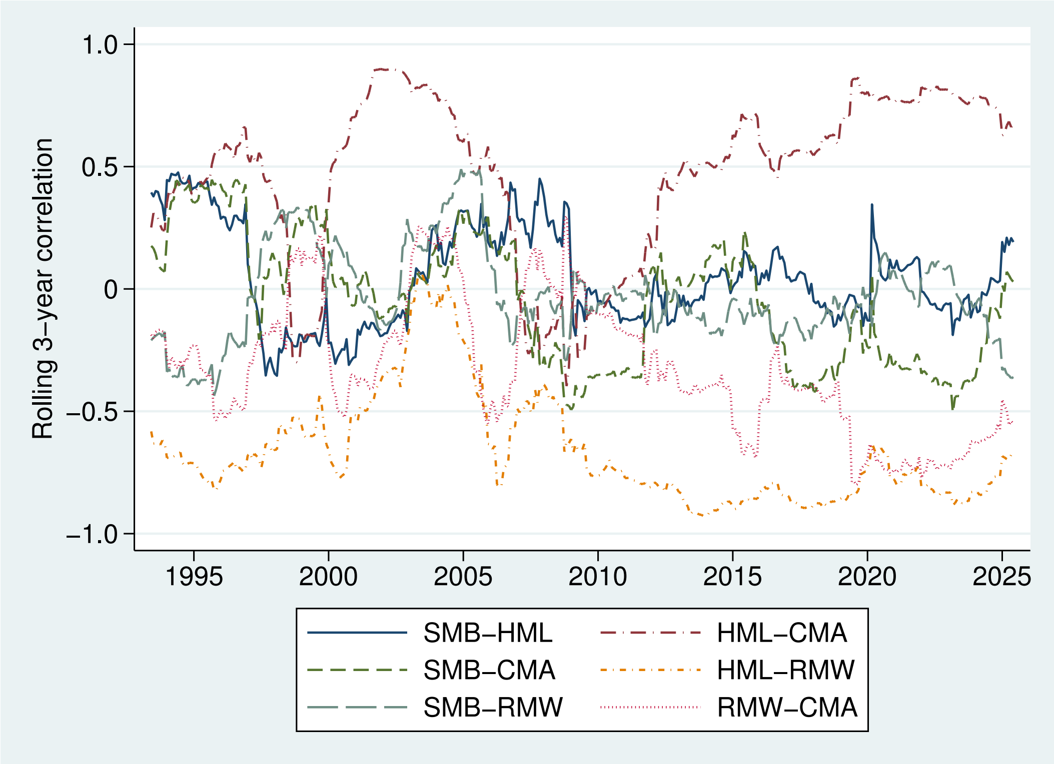

Fig. 2.1. This figure shows the 3-year rolling correlation of European non-market factors proposed in Fama and French (2015). SMB denotes the size factor, HML the value factor, RMW the profitability factor, and CMA the investment factor. Factor returns from July 1991 to June 2025 are obtained from Kenneth French's data library.

included in the well-known European Monetary Union (EMU) stock market benchmark from MSCI and comprises on average 913 firms each month from June 1991 to June 2025 (378,281 firm-month observations). Because our main analysis in this study includes the investigation of potential relations of CF and DR news with macroeconomic variables, this sample selection benefits from the conceptual advantage that EMU countries have established a single monetary policy. This makes the interpretation of our macroeconomic analysis straightforward.

Our main findings are as follows: In line with Lochstoer and Tetlock (2020) for the U.S., cash flow shocks tend to dominate the returns of portfolios representing the size, value, investment, profitability, and momentum anomaly. Their contribution to the return variance of these anomaly portfolios is within the range from 66.64% for the value anomaly, and 82.65% for the size anomaly. In contrast, discount rate shocks contribute less than 16.54% (momentum) among all anomalies. On average, cash flow shocks also explain 89.71% of market-adjusted return variance at the firm-level, while discount rate shocks are a less important component. Interestingly, while some of the cash flow and discount rate shocks are correlated with shocks to macroeconomic variables, correlations are at best modest in magnitudes, ranging from only -0.22 to 0.15. In summary, we find that anomaly portfolios tend to have a common component located in their cash flow news, but this unobserved component seems to be uncorrelated with shocks in standard macroeconomic variables. Applying a time-varying vector auto-regression (TVP–VAR) based extended joint connectedness framework proposed in Balcilar et al. (2021) for these CF and DR shocks, we find that, on average, 40.22% of a shock in one of anomaly CF or DR news spills over to all others. The largest shocks are obtained within discount rates, whereas their spillovers to the CF shock category seems to be modest. Despite the dominance of CF shocks among anomaly portfolio returns, DR shocks seem to play a more important role in carrying shocks from one anomaly to the others. These findings, overall, provide guidance for the development of asset pricing theories, as shown in Lochstoer and Tetlock (2020).

The remainder of this study is organized as follows. Section 2.2 describes the data and variables used in this paper. Section 2.3 introduces the return-decomposition model and presents estimation results. Section 2.4.1 analyzes the importance of CF and DR shocks for the variation of anomaly portfolio returns. Section 2.4.2 analyzes the relation between estimated CF and DR shocks with shocks in macroeconomic aggregates. Section 2.5 presents results for the connectedness approach proposed in Balcilar et al. (2021). Finally, Section 2.6 concludes.

2.2 Data description

We study an integrated European stock market sample that consists of firms from ten developed countries of the European Monetary Union (EMU): Austria, Belgium, Finland, France, Germany, Ireland, Italy, Netherlands, Portugal, and Spain. Our sample selection resembles the countries included in the well-known EMU stock market benchmark from MSCI and comprises on average 913 firms each month from June 1991 to June 2025 (378,281 firm-month observations). Since these members of the euro area have established a single monetary policy, the conceptual advantage of this sample is that the interpretation of our presented macroeconomic analysis is straightforward.

We collect monthly total return data on common stocks from LSEG Eikon/Datastream and firm-level accounting information from Worldscope. To ensure that accounting information is known before the returns are calculated, we match the latest accounting information for the fiscal year ending in the previous calendar year with stock returns from July of the current year to June of the following year throughout the paper (see Fama and French (1992) and Fama and French (1993)). All data are denominated in euros. We obtain European factor returns for the five-factor model proposed in Fama and French (2015) from Kenneth French's data library. Since these European factor returns are calculated in U.S.-$, we converted them into Euro following the procedure outlined in Glück et al. (2021). The risk-free rate is proxied by FIBOR before 1999 and EURIBOR thereafter.

We follow Ince and Porter (2006), Griffin et al. (2010), and Schmidt et al. (2017) and apply the suggested static and dynamic screens to ensure a high data quality and to exclude non-common equity securities like ADRs, Funds, or REITS. We delete monthly returns greater than 990% or abnormal high returns followed by strong reversal as in Ince and Porter (2006).[1] According to Fama and French (1992) we exclude financial firms indicated by an SIC code between 6,000 and 6,999 or an ICB industry number of 30 or 35. If the SIC code or ICB industry number is missing, we disregard the firm. We also exclude firm observations with negative book equity. To ensure that our results are not driven by very small and illiquid firms, we follow Ang et al. (2009) and eliminate the 5% of firms with lowest market capitalization measured at the end of June in each country from our sample.

All variables used in this study are winsorized at the 0.5%/99.5% levels and defined based on Lochstoer and Tetlock (2020) as follows. We compute monthly log stock excess returns *lnRETEXC* by subtracting the log change of the risk-free rate from the log nominal stock return. Similar with Gerakos and Linnainmaa (2018), *lnME* is the three-year change in log market equity (stock price multiplied by the number of shares outstanding), rather than the level of log market equity, to ensure stationarity in our VAR system. *lnBM* is the log of book-to-market, which is the ratio of book equity to market equity for the fiscal year ending in the previous calendar year. According to Cooper et al. (2008) and Fama and French (2015), we proxy for investment using the three-year average (log) growth in total assets as our long-horizon investment characteristic *lnINV*. Following Fama and French (2015), we construct profitability as revenues minus cost of goods sold and interest expense, all divided by book equity.[2] We use the log of one plus profitability in the VAR (*lnPROF*). The log of cumulative prior twelve-month stock return (Jegadeesh and Titman

[1] If r_{t-1} or r_t is greater than 300% per month and $(1 + r_{t-1}) \cdot (1 + r_t) - 1 < 0.5$, we eliminate both return observations.

[2] As in Walkshäusl (2021), we do not include selling, general, and administrative expenses because this data item is not broadly available among non-U.S. firms.

(1993)), skipping the most recent month, is denoted as momentum (*lnMOM*). *lnROE* is the log of income before extraordinary items divided by book value of equity. In general, we transform each variable by adding one and taking its log, except for book-to-market and change in size (*lnME*).

Table 2.1: Descriptive statistics.

This table shows summary statistics for the countries included in the European sample and the variables used in this study. Panel A reports the average number of firms in each country and the country's average percentage weight in terms of total market equity. Panel B reports time-series averages of the cross-sectional mean, standard deviation, 25th percentile, median, and 75th percentile of the variables for the sample period from June 1991 to June 2025. *lnRETEXC* denotes the monthly stock return (in %) in excess of the risk-free rate (FIBOR before 1999 and EURIBOR thereafter). *lnME* is the three-year change in market equity (stock price multiplied by the number of shares outstanding). *lnBM* is the book value of equity at the fiscal year-end divided by the market value of equity at the end of December. *lnINV* is the three-year average growth in total assets. *lnPROF* is revenues minus cost of goods sold and interest expense, all divided by book value of equity. *lnMOM* is the cumulative prior twelve-month stock return skipping the most recent month. *lnROE* is income before extraordinary items divided by book value of equity. We measure all variables in natural logs, adding one before taking the log except for book-to-market and change in size (*lnME*). See the text for details.

Panel A: Developed sample countries

Country	Firms	Weight	Country	Firms	Weight
Austria	25	1.11	Ireland	19	1.01
Belgium	39	3.83	Italy	122	8.59
Finland	70	4.16	Netherlands	67	10.05
France	270	35.35	Portugal	25	1.05
Germany	228	26.38	Spain	57	8.46

Panel B: Variables

Variable	Mean	St. Dev.	25th	Median	75th
$lnRETEXC$	−0.01	10.44	−4.83	−0.10	4.83
$lnME$	0.12	0.67	−0.25	0.12	0.50
$lnBM$	−0.37	0.63	−0.79	−0.39	0.02
$lnINV$	0.22	0.39	0.01	0.16	0.36
$lnPROF$	0.52	0.36	0.27	0.45	0.68
$lnMOM$	0.03	0.30	−0.13	0.04	0.20
$lnROE$	0.06	0.21	0.02	0.09	0.15

Panel A of Table 2.1 shows the average number of firm observations per country during the sample period from June 1991 to June 2025. In line with their economic importance in the European Monetary Union, the three largest stock markets in the sample, being France, Germany, and Italy, contribute by far the largest portion of observations. France accounts on average for 270 firms and 35.35% of the sample's total market equity, whereas

Germany contributes 228 firms and 26.38% of total market equity. Together, they comprise 61.73% of total market equity during the sample period.

Panel B of Table 2.1 shows summary statistics for the variables used in this study. The average stock's excess return is close to zero, implying that the typical firm is not able to outperform the risk-free rate as documented in Bessembinder (2018). The average firm has a log book-to-market ratio of -0.37, which implies a market-to-book ratio of $e^{0.37} = 1.45$, and generates a profit of approx. 6% its book value. On average, a typical firm in our EMU sample has a market capitalization of €2739.08 million (untabulated), however, most of our observed firms are much smaller as indicated by the median market capitalization of €247.17 million.

2.3 Decomposing stock returns

In this section, we specify a VAR system with panel regressions for firm-specific variables and time-series regressions for aggregate variables at a monthly frequency, following the approach proposed in Lochstoer and Tetlock (2020). We first introduce the VAR specification in Section 2.3.1 and present coefficient estimates and implications of the VAR system in Section 2.3.2.

2.3.1 VAR estimation model

We closely follow Lochstoer and Tetlock (2020) in estimating shocks to firm-level stock returns that can be decomposed into cash flow (CF) and discount rates (DR) shocks. To allow predictive estimation coefficients to differ across firms and over time as in Vuolteenaho (2002) and Cohen et al. (2003), two separate VAR(1) systems are implemented.

To begin with, we first consider the aggregated VAR^{agg} model

$$Z_{t+1} = \mu^{agg} + A^{agg} Z_t + \epsilon_{t+1}^{agg}, \tag{2.1}$$

where $Z_t = [r_t^{agg}; X_t^{agg}]$ is a vector of the value-weighted, average excess return r_t^{agg} as its first element, followed by aggregated firm characteristics X_t^{agg}, such as the value-weighted average book-to-market ratio at time t. We use value-weighted averages for all variables to aggregate firm-level data. As outlined in Campbell (1991), aggregated discount rate shocks are directly computed by

$$\begin{aligned} DR^{agg}_{t+1} &= E_{t+1}\sum_{j=2}^{\infty}\kappa^{j-1}r^{agg}_{t+j} - E_t\sum_{j=2}^{\infty}\kappa^{j-1}r^{agg}_{t+j} \\ &= e_1^{'}\kappa A^{agg}(I_{K^{agg}} - \kappa A^{agg})^{-1}\epsilon^{agg}_{t+1}. \end{aligned} \tag{2.2}$$

e_1 is a vector with one as its first element and zeros elsewhere, $I_{K^{agg}}$ is an identity matrix, and $\kappa < 1$ is the inverse of the gross discount rate.[3] Because of the shorter perspective of our monthly analysis, we set $\kappa = 0.9999$.[4]

Second, for the cross-section of firm-level returns, we estimate a market-adjusted panel system (VAR^{ma})

$$Z_{i,t+1} = \mu^{ma} + A^{ma}Z_{i,t} + \epsilon^{ma}_{i,t+1}, \tag{2.3}$$

where $Z_{i,t} = \left[r^{ma}_{i,t}; X^{ma}_{i,t}\right]$ is a vector of market-adjusted variables with the market-adjusted log return $r^{ma}_{i,t} \equiv r_{i,t} - r^{agg}_t$ as its first element. Firm-level market-adjusted discount rate shocks are then obtained by

$$DR^{ma}_{i,t+1} = i_1^{'}\kappa A^{ma}(I_{K^{ma}} - \kappa A^{ma})^{-1}\epsilon^{ma}_{i,t+1}, \tag{2.4}$$

where i_1 is a vector with one as its first element and zeros elsewhere, and $I_{K^{ma}}$ an identity matrix. To control for possible country effects, Equation (2.4) is estimated with country dummies. As the number of observable firms in our sample increases over time, the estimated coefficients would be dominated by the larger cross-sections in most recent year. To ensure that each month is weighted equally, we apply a weight $w_t = 1/N_t$ to each firm-month observation, where N_t is the total number of firms existing in month t.

The present-value identity proposed in Campbell (1991) emphasizes that we can decompose shocks to log stock returns into shocks to expectation of cash flows and returns, so based on our VAR systems presented in Equations (2.2) and (2.4), we follow Lochstoer and Tetlock (2020) and obtain cash flow shocks by

$$\begin{aligned} CF^{agg}_{t+1} &= r^{agg}_{t+1} - E_t\left[r^{agg}_{t+1}\right] + DR^{agg}_{t+1} \\ &= e_1^{'}\Big(I_{K^{agg}} + \kappa A^{agg}(I_{K^{agg}} - \kappa A^{agg})^{-1}\Big)\epsilon^{agg}_{t+1}, \end{aligned} \tag{2.5}$$

[3]According to Campbell (1991) and Campbell and Shiller (1988b), κ denotes the average log dividend yield. Another interpretation is provided by the framework proposed in Campbell (1993) and Campbell (1996), that define $\kappa \equiv 1 - e^{c-w}$, where $c - w$ is the mean log consumption-wealth ratio.

[4]Lochstoer and Tetlock (2020) use $\kappa = 0.95$ for an annual perspective, and a value of 0.99 is typically used for a quarterly period.

$$
\begin{aligned}
CF^{ma}_{i,t+1} &= r^{ma}_{i,t+1} - E_t\left[r^{ma}_{i,t+1}\right] + DR^{ma}_{i,t+1} \\
&= i_1^{'}\Big(I_{K^{ma}} + \kappa A^{ma}(I_{K^{ma}} - \kappa A^{ma})^{-1}\Big)\epsilon^{ma}_{i,t+1}.
\end{aligned} \tag{2.6}
$$

Finally, firm-level total DR and CF shocks are the sum of the aggregated and market-adjusted components:

$$DR_{i,t} = DR^{agg}_t + DR^{ma}_{i,t}, \tag{2.7}$$

$$CF_{i,t} = CF^{agg}_t + CF^{ma}_{i,t}. \tag{2.8}$$

This approach allows the predictive coefficients in the VAR to differ across firms (VAR^{ma}) and over time (VAR^{agg}), which is an important property to match the data.

Following Vuolteenaho (2002) and Lochstoer and Tetlock (2020), we analyze pseudo-firms, which comprise portfolios with a 10% weight in the risk-free rate and a 90% position in the firms' stocks. The reason for this is that if a firm goes bankrupt and its stock price equals zero, its gross return is zero, which means that its log return is undefined. We similarly adjust the pseudo-firms' book-to-market ratio, profitability, and other firm characteristics accordingly.

As in Lochstoer and Tetlock (2020), we analyze CF and DR shocks to five long-short anomaly portfolios for our EMU sample. Each of these portfolios takes long (short) positions in the top (bottom) quintile of stocks sorted by one of the five anomaly characteristics. We construct the CF and DR shocks to the long and short portfolios by value-weighting the CF and DR shocks to the firms in these portfolios. We then compute the long-short portfolios' CF and DR shocks as the difference between the long and short legs.

2.3.2 VAR estimation results

For estimating the market-adjusted panel system VAR^{ma} as outlined in the previous section, we adjust all firm-specific variables by subtracting the corresponding market-level variables. The aggregated (market) variables are value-weighted averages of the unadjusted versions of these variables, which are used in the estimation of the aggregated system VAR^{agg}.

Both VAR systems include seven variables: Monthly log excess returns ($lnRETEXC$), the five anomaly characteristics, namely, *lnME*, *lnBM*, *lnINV*, *lnPROF*, and *lnMOM*, and *lnROE*. As in Lochstoer and Tetlock (2020), we restrict the coefficients on the lagged *lnROE* variables to be zero in the coefficient matrix, since lagged *lnROE* does not add predictive power beyond the other variables already included in the VAR. Overall, the VAR specifications are clearly motivated by the models proposed in Fama and French (2015) and Carhart (1997).

Table 2.2 reports estimates of the predictive coefficients A^{agg} from the aggregated VAR system in Panel A, and estimates of the predictive coefficients A^{ma} in the market-adjusted panel VAR in Panel B.

The first column of Panel A shows the forecasting regression for aggregated log one-month (market) returns. Stock market returns are predictable in the long-run (see e.g., Cochrane (2011)), but this is not the case in our short-run VAR forecasts. Most coefficients for lagged variables have signs that are expected, e.g., positive in case of book-to-market ratio, but are statistically insignificant. The only exception is asset growth ($lnINV$) which is a highly significant (and as expected negative) -0.064. However, the p-value for the F-test of joint significance of all coefficients is a significant 0.01.

The very weakly pronounced predictability of one-month market returns is further indicated by the low value of 0.03 for the adj. R^2. However, R^2 values are a poor measure for the economic implications of market predictability as outlined in Cochrane (2011). A better metric is the volatility of expected, future market returns $\sigma\left(E_t\left[lnRETEXC_{t+1}\right]\right)$ which is a small 0.97%. Again, this low value implies that it is hard to predict the market return in the short-run, but considering the mean of expected returns, 0.26%, the time-variation exceeds its mean more than three times, which seems to be in line with the longstanding equity premium puzzle (see Mehra and Prescott (1985)).

The remaining columns in Panel A of Table 2.2 show the forecasting regressions for the aggregate predictors. The most persistent predictors are the aggregated *lnBM* which has a persistence coefficient of 0.980 and the aggregated *lnME* (0.973), thus being the primary determinant of our short-run (aggregated) predictability. This is close to their autocorrelation coefficient of 0.972 for *lnBM* and 0.977 for *lnME*, indicating that the presence of the other regressors does not substantially increase the persistence coefficients for book-to-market and asset growth.

Coefficient estimates for our market-adjusted panel VAR system are reported in Panel B of Table 2.2. Again, looking at the first column for the market-adjusted return regressions reveals that most coefficients are statistically significant, as documented in the vast

Table 2.2: Aggregated and market-adjusted VAR estimates.

The table reports results from estimating the aggregated VAR (Panel A) and the market-adjusted panel VAR (Panel B) as outlined in Section 2.3.1. The variables used are: Log monthly stock excess returns (*lnRETEXC*), three-year change in log market equity (*lnME*), log book-to-market ratio (*lnBM*), log three-year growth in total assets (*lnINV*), log profitability (*lnPROF*), and log cumulative past twelve-month return, skipping the most recent month (*lnMOM*). We also include log return on equity (*lnROE*) as a dependent variable in the VAR systems, but restrict all coefficients on this variable's lag to zero. All VAR systems additionally include country dummies. The sample period is June 1991 to June 2025. Robust t-statistics based on clustered standard errors (clustered by month and firm) are provided in parenthesis. */**/*** indicate significance at the 10%/5%/1% level. The R^2 value is adjusted for degrees of freedom. The aggregated VAR system considers $N = 408$ observations and the market-adjusted panel VAR considers $N = 364,767$ observations.

	Dependent Variables						
	$lnRETEXC_t$	$lnME_t$	$lnBM_t$	$lnINV_t$	$lnPROF_t$	$lnMOM_t$	$lnROE_t$
Panel A: Aggregated VAR							
$lnRETEXC_{t-1}$	0.064	0.002	0.029	−0.028	0.002	0.878	−0.017*
	(1.03)	(0.02)	(0.61)	(−0.93)	(0.12)	(>10)	(−1.82)
$lnME_{t-1}$	−0.003	0.973***	−0.009	0.020**	0.005	0.009	0.006**
	(−0.40)	(>10)	(−1.15)	(2.34)	(0.91)	(0.89)	(2.07)
$lnBM_{t-1}$	0.012	0.005	0.980***	−0.002	−0.003	0.023*	−0.003
	(1.07)	(0.26)	(>10)	(−0.23)	(−0.67)	(1.75)	(−0.97)
$lnINV_{t-1}$	(−0.064)***	−0.102***	0.004	0.962***	0.003	0.008	−0.009
	(−3.10)	(−2.85)	(0.19)	(>10)	(0.28)	(0.4)	(−1.50)
$lnPROF_{t-1}$	0.042	0.027	0.097	−0.063	0.877***	0.063	−0.026
	(0.77)	(0.29)	(1.59)	(−1.18)	(>10)	(1.07)	(−1.12)
$lnMOM_{t-1}$	−0.006	0.029	−0.039*	−0.011	−0.012	0.893***	−0.007
	(−0.30)	(1.02)	(−1.73)	(−0.87)	(−1.02)	(>10)	(−1.07)
Adj. R^2	0.03	0.96	0.96	0.94	0.78	0.92	0.88
Panel B: Market-adjusted panel VAR							
$lnRETEXC_{t-1}$	−0.014*	0.014*	0.014*	−0.014***	0.002	0.794***	0.005**
	(−1.87)	(1.69)	(1.85)	(−2.84)	(0.75)	(>10)	(2.40)
$lnME_{t-1}$	−0.003*	0.960***	−0.009***	0.016***	−0.001*	0.003*	0.003***
	(−1.89)	(>10)	(−3.97)	(4.39)	(−1.87)	(1.83)	(5.15)
$lnBM_{t-1}$	0.003***	0.002*	0.988***	−0.001*	−0.001	0.007***	−0.002***
	(3.94)	(1.91)	(>10)	(−1.70)	(−1.59)	(7.71)	(−4.43)
$lnINV_{t-1}$	−0.004***	−0.013***	0.002	0.963***	0.002**	0.001	−0.003***
	(−2.77)	(−5.35)	(1.17)	(>10)	(2.51)	(0.52)	(−4.53)
$lnPROF_{t-1}$	0.001	0.003*	−0.002	0.000	0.991***	0.003***	0.002***
	(1.21)	(1.83)	(−1.45)	(−0.15)	(>10)	(3.84)	(2.99)
$lnMOM_{t-1}$	0.025***	0.073***	−0.030***	−0.014***	0.001	0.904***	0.011***
	(7.77)	(>10)	(−5.02)	(−3.59)	(1.16)	(>10)	(5.53)
Adj. R^2	0.01	0.95	0.98	0.96	0.98	0.92	0.93

literature for the cross-section of stock returns (see e.g., Harvey and Liu (2016)).

We find an economically strong effect for the book-to-market ratio, asset growth, and momentum. As a gauge of magnitude, a one standard deviation increase in momentum (i.e., 0.30, see Table 2.1) corresponds to an increase in one-month ahead excess returns of 0.75%, or 9% annualized. We find that lagged returns are negatively related with *lnRETEXC*, but given its merely significance, our results indicate that the reversal effect as documented in Jegadeesh (1990) is only weakly pronounced in our EMU sample. The same holds for three-year change in market equity (*lnME*), and we further find that profitability does not add to the predictability of our short-run, market-adjusted panel VAR system. The modest adj. R^2 of 0.01 is quite typical for firm-level stock return forecasts. The diagonal elements of A^{ma} show the persistence of each predictive variable. We find very high persistence coefficients larger than 0.96 for *lnPROF*, *lnBM*, *lnINV*, and *lnME*, while momentum has a smaller persistence coefficient of only 0.904. Besides market-adjusted returns, however, all persistence coefficients are very well measured as shown by astonishing huge t-statistics exceeding the value of ten in all cases.

2.4 Anomaly return decomposition and the macroeconomy

2.4.1 Firm and anomaly return analysis

We now extend the results of Lochstoer and Tetlock (2020) for our EMU sample and examine whether cash flow shocks, discount rate shocks, or their according covariance is the main driver of monthly stock returns. We use the separately estimated aggregated and market-adjusted CF and DR components as described in Section 2.3. A firms' total excess-return shock is obtained by adding the respective components of market-adjusted and aggregated shocks.

We also analyze CF and DR shocks at the anomaly portfolio level for the size-, value-, investment-, profitability-, and momentum-effect. As emphasized in Lochstoer and Tetlock (2020), only correlated shocks to firms remain when aggregating firm-level shocks. For that reason, results for the portfolio return variance decomposition can be very different from the firm return variance decomposition.

Whereas the VAR system uses log-transformed variables, or variants like three-year change in market equity (*lnME*) as a proxy for size to ensure stationarity, the portfolio sorts in this section are based on traditional variable definitions as in Fama and French (2015) to be consistent with other empirical studies on anomalies. For that reason, each June, we form quintile portfolios by sorting stocks based on size (i.e., stock price multiplied by the number of shares outstanding as of June), value (ratio of book equity to market equity for the fiscal year ending in the previous calendar year), investment (annual change in total assets divided by lagged total assets), and profitability (revenues minus cost of goods sold and interest expense, all divided by book equity). Quintile portfolios based on momentum (cumulative prior twelve month stock return, skipping the most recent month) are updated each month.

Table 2.3: Return variance decomposition.

This table shows the variance decomposition of firm-, market-, and portfolio-level excess returns into cash flow and discount rate shocks based on the market-adjusted panel VAR(1) and the aggregated VAR(1) system outlined in Section 2.3. Panel A decomposes market-adjusted log firm-level returns (first row) and the according value-weighted average of log firm-level returns denoted as market return (row three). The decomposition of firm returns (row two) is based on combining components of firm market-adjusted returns and market returns. Panel B decomposes long-short anomaly portfolio returns into cash flow and discount rate shocks. The portfolio returns are the difference between the top quintile portfolio and the bottom quintile portfolio, and the portfolios are based on sorts at the end of each June, except for momentum, which is monthly updated. The sample spans the period from June 1991 to June 2025. Standard errors appear in parenthesis and account for estimation uncertainty from sampling variation and from estimating the VAR coefficients, as well as for heteroskedasticity and contemporaneous cross-correlation of residuals.

	Fraction of Portfolio Return Variance			
	$var(DR)$	$var(CF)$	$-2cov(DR, CF)$	$cor(DR, CF)$
Panel A: Stock return components				
Firm market-adjusted return	9.16%	89.71%	1.12%	0.02
	(8.40%)	(11.23%)	(11.95%)	(0.21)
Firm return	15.53%	80.54%	3.39%	0.06
	(3.46%)	(6.73%)	(12.17%)	(0.21)
Market return	49.66%	41.28%	9.06%	0.10
	(4.77%)	(8.09%)	(20.37%)	(0.26)
Panel B: Anomalies				
Size	12.42%	82.65%	4.93%	0.08
	(7.58%)	(8.90%)	(2.66%)	(0.07)
Book-to-market	15.10%	66.64%	18.26%	0.29
	(8.16%)	(8.97%)	(2.71%)	(0.07)
Investment	13.54%	77.06%	9.40%	0.15
	(7.44%)	(8.96%)	(2.67%)	(0.07)
Profitability	14.46%	78.51%	7.03%	0.10
	(8.87%)	(8.91%)	(2.67%)	(0.07)
Momentum	16.54%	70.38%	13.07%	0.19
	(6.94%)	(9.01%)	(2.69%)	(0.07)

Table 2.3 shows the decomposition of log return variance into DR and CF components. Standard errors, calculated as in Lochstoer and Tetlock (2020), are reported in parentheses and account for estimation uncertainty from sampling variation and from estimating the

VAR coefficients, as well as for heteroskedasticity and contemporaneous cross-correlation of residuals.

First, considering the firm-level return decomposition in Panel A of Table 2.3, we observe that CF shocks are able to explain 89.71% of the market-adjusted return variance, while DR shocks are a less important component. In contrast, the market return is strongly impacted by DR variation, a well-known fact as summarized in Cochrane (2008) and Cochrane (2011). Both market-adjusted and market components add to the total firm return, thus we find less extreme results for their decomposition. However, CF news also explain the majority (80.54%) of the total firm return variance. Overall, these findings for the European Monetary Union are as expected from the literature (see e.g., Vuolteenaho (2002)).

Interestingly, the correlation between CF and DR shocks is positive for market-adjusted and total firm returns. Previous U.S. findings document this behavior to emerge only among microcaps, which are excluded prior to our analysis. In general, a negative relation economically implies investor overreaction as proposed by some studies. This overreaction could arise for behavioral (see e.g., positive feedback trading and extrapolation of cash flows in Barberis et al. (1998) and Hong and Stein (1999)) or rational reasons (see e.g., CF shock induced reduction of exposure to systematic risk as in Babenko et al. (2016)). However, we may not see these models' economic implications in our short-run VAR analysis, which takes the predictive horizon of just one month.

Panel B of Table 2.3 shows the decomposition of the return variance for anomaly portfolios proposed in Fama and French (2015) and Carhart (1997). We find that the contribution of CF shocks to the return variance of anomaly portfolios is within the range from 66.64% for the value anomaly and 82.65% for the size anomaly. In contrast, DR shocks contribute less than 16.54% (momentum) in all five anomalies. While this finding confirms the results of Lochstoer and Tetlock (2020) outside the U.S., the dominance of CF shocks among anomaly portfolios is still puzzling.

So far, we have analyzed individual anomalies and found that CF news dominate their returns, so the key question arises: Is there a common component in CF news across the anomalies? Looking at the (ex-post) mean-variance efficient (MVE) portfolio comprising all anomalies without the market factor, we observe that it achieves an annualized Sharpe ratio of 0.99, more than two times the Sharpe ratio of the European market portfolio (0.45). If we decompose its returns similar to the anomaly portfolios, we notice that

89.05% of its variance stems from CF news.[5] These findings are important, because shocks to the return of the MVE portfolio are proportional to shocks to the unobservable stochastic discount factor (SDF) and may reflect risks associated with the marginal utility of investors. Most researchers and practitioners put the market portfolio in the center of the risk-return space, which is additionally spanned by risk factors related with firm characteristics, where investors may tilt away from (see e.g., Fama and French (2004)). However, Kolokolova et al. (2021) find that most equity indices around the world are not efficient under the stochastic dominance criteria and investors could improve their expected utility by alternative asset allocations.[6] Our return decomposition approach is able to provide a numerical order of magnitude: If we regress the MVE portfolio returns on the European market excess returns, we observe a negligible $adj.R^2$ of 1.97%, thus nearly all variation in the implied SDF is attributable to non-market risk-factors.

2.4.2 Anomaly cash flow and discount rate shocks and the macroeconomy

Cash flow risk and discount rate risk are often tied to measures of macroeconomic activity or proxies for time-varying risk aversion. Take e.g., the value anomaly. Campbell and Vuolteenaho (2004) and Da and Warachka (2009) explain the value premium by analyzing cash flow betas and find that value stocks have higher cash flow betas than growth stocks. Golubov and Konstantinidi (2019) show that the value premium may represent compensation for aggregate cash flow risk. In addition, Asness et al. (2013) provide evidence that value and/or momentum are related with consumption growth, as also documented in Bansal and Yaron (2004), Parker and Julliard (2005), or Malloy et al. (2009).

In this sense, we relate CF and DR shocks of our anomaly portfolio returns to shocks of macroeconomic activities in this section. We estimate each aggregate shock as the residual from a first-order autoregressive (AR(1)) model of the relevant time series.

Some aggregate shocks reflect clear macroeconomic cash flow shocks like e.g., consumption and GDP growth. Other represent undisputed discount rate shocks like the term and default spread, or have no clear assignment to these categories, e.g., export growth.

[5] If we calculate the MVE portfolio including the market factor, the Sharpe ratio remains marginally increases to 1.0 (ann.). Still, 85.09% of return variance is attributable to CF news.

[6] There is plenty of empirical evidence. DeMiguel et al. (2009), Novy-Marx (2013), and Brière and Szafarz (2021) show that the market portfolio is clearly dominated when forming mean-variance efficient portfolios comprising other anomalies. Harvey and Liu (2021) even state "the market factor has obvious theoretical appeal, yet its empirical success is limited" (p. 414).

However, their conceptual advantage is to be observable at a monthly frequency, matching the predictive horizon of our VAR system.[7] Macroeconomic data is retrieved by LSEG Eikon and defined as follows. Inflation denotes the one-month change in the HICP index including all items (EMCPHARMF). We further consider the change in unemployment in euro zone countries (EMUNPTOTO), change in M2 money supply (EMM2....B), change in the term spread which is the difference in 10-Year (EMGBOND.) and 2-Year (EMECB2Y.) EMU government bond yields, change in the European consumer confidence indicator (EKCNFCONQ, since 1999),[8] change in extra-EMU exports (EMEXPGDSB), change in industrial production (EKIPTOT.G), change in the sentiment index proposed in Baker and Wurgler (2006), change in the default spread[9] (difference in EMU corporate bond yields (IBCRPAL) and 10-Year EMU government bond yields after 1999, and difference in BAA-rated and AAA-rated U.S.-corporate bond yields prior to 1999, taken from FRED), change in GDP (EMGDP...D), and change in consumption[10] (EMCNPER.D).

For ease of interpretation, we multiply the long-short returns of the investment and size portfolios by -1 before computing correlations, to reflect the fact that the strategies' long-leg is based on small characteristics (i.e., small market capitalization, resp., conservative investment behavior).

[7]The only exceptions are GDP and consumption, which are stated in quarterly figures. We cumulate CF and DR shocks within each calendar quarter in order to calculate their correlations with GDP and consumption.

[8]The consumer confidence indicator is the arithmetic average of the balances (in percentage points) of the answers to the questions on the financial situation of households, the general economic situation, unemployment expectations (with inverted sign) and savings, all over the next 12 months.

[9]EMU corporate bond yields are not available prior to 1999. The correlation between the U.S. default spread and our EMU default spread since 1999 exceeds 0.70., so we consider the U.S. default spread to be appropriate for earlier periods.

[10]The value of the consumption goods and services acquired by households, whether by purchase in general, or by transfer from government units or NPISHs.

Table 2.4: Correlations of cash flow and discount rate news with macroeconomic metrics.

This table reports correlations of the market, the anomaly portfolios', and the mean-variance efficient combinations' (MVE ex market), cash flow and discount rate shocks with shocks to several aggregate macroeconomic metrics: One-month change in inflation (HICP all items), change in unemployment growth, change in M2 money supply, change in term spread (difference in 10-Year and 2-Year EMU government bond yields), change in consumer confidence indicator, change in exports (extra-EMU), change in industrial production, change in the sentiment index proposed in Baker and Wurgler (2006), change in default spread (difference in EMU corporate bond yield and 10-Year EMU government bond yield after 1999 and difference in BAA-rated and AAA-rated U.S.-corporate bond yields before), change in GDP, and change in consumption. All macroeconomic variables are measured at a monthly frequency with the exception of GDP and consumption which are quarterly measured. For ease of interpretation, we multiply the long-short returns of the investment and size portfolios by −1 before computing correlations, so that these portfolios have positive premiums. The sample period is June 1991 to June 2025. */**/*** indicate significance at the 10%/5%/1% level based on Fisher's z-transformation (Fisher (1915)).

	Dependent Variables										
	Inflation	Unempl.	M2	Term spread	Consumer Conf.	Export	Indust. Prod.	Sentiment	Default spread	GDP	Consumption
Panel A: Cash flow shock correlations											
Market	−0.01	0.00	−0.06	0.01	0.06	−0.04	0.02	0.01	−0.22***	0.03	0.02
(-) Size	−0.08	0.04	−0.05	0.03	−0.01	−0.07	−0.02	0.11	0.00	0.08	0.03
Book-to-market	0.02	0.05	−0.11**	0.01	0.15***	0.07	0.07	0.11**	−0.04	0.06	−0.01
(-) Investment	0.04	0.08*	−0.09*	0.04	0.01	0.07	0.05	0.10*	0.05	−0.05	−0.07
Profitability	−0.03	−0.01	−0.02	0.04	0.10**	0.12**	0.08	0.03	−0.06	0.04	0.05
Momentum	0.08	−0.01	0.10**	0.08	−0.12**	0.00	0.02	0.06	0.11**	0.05	0.03
MVE (ex market)	0.00	0.03	0.00	0.06	−0.02	−0.04	0.02	0.13**	0.04	0.09	0.02
Panel B: Discount rate shock correlations											
Market	0.02	0.01	0.00	0.08	−0.15***	−0.10**	−0.08	0.04	0.19	−0.04	0.01
(-) Size	0.00	−0.02	−0.05	0.00	0.04	0.02	0.04	0.01	−0.03	−0.13	−0.13
Book-to-market	0.00	−0.02	−0.07	0.02	0.02	0.01	0.01	0.05	−0.01	−0.07	−0.05
(-) Investment	0.00	0.04	−0.07	−0.05	0.03	−0.04	−0.01	−0.03	−0.07	−0.06	−0.03
Profitability	−0.01	−0.01	−0.04	0.00	−0.03	−0.01	−0.05	−0.04	−0.03	0.06	0.02
Momentum	0.04	0.05	0.01	−0.02	0.01	0.01	0.01	−0.11**	−0.04	0.08	0.08
MVE (ex market)	0.04	0.01	−0.05	0.00	0.05	0.04	0.05	−0.05	−0.05	−0.06	−0.05

Table 2.4 reports correlations of CF and DR shocks to the market, the mean-variance efficient portfolio (ex market), and anomaly portfolios with aggregate macroeconomic shocks. As in Lochstoer and Tetlock (2020), standard errors for correlation estimates consider sampling variation based on Fisher's z-transformation (Fisher (1915)).

Market CF shocks are negatively correlated with shocks to the default spread, which is a plausible measure of risk aversion or discount rate shocks. On the other hand, we observe a positive correlation of market DR shocks with the default spread, consistent with the latter being a measure when risk aversion is high. Similarly, market DR shocks are negatively correlated with consumer confidence and export shocks, implying that market discount rates increase in recessions. These findings are in line with Lochstoer and Tetlock (2020) and earlier literature.

With respect to anomaly return shocks, we observe several interesting patterns in their correlations with macroeconomic shocks. We do not observe significant correlations of CF or DR shocks between the MVE (ex market) portfolio and our macroeconomic variables, even for clear proxies for risk aversion like default spread or term spread. The only exception is investor sentiment, which is positively related with CF shocks which is quite intuitive. Further, the value anomaly shows positive correlations of CF shocks with shocks in sentiment and consumer confidence, and a negative correlation with M2 money supply. While some of the CF and DR correlations seem to be significant, they are at best modest in magnitudes, ranging from -0.22 to 0.15. Interestingly, consumer confidence shocks are correlated with multiple anomaly CF shocks. We observe a positive relation with book-to-market (0.15) and profitability (0.10), whereas the correlation is a negative -0.12 with momentum.

Overall, these results have important implications for asset pricing theories. We have documented that CF news dominate firm returns, and variation in CF news has a commonality across anomaly portfolios. Their characteristics are, however, proxies for different exposures to this common CF news component, but this unobserved factor seems to be uncorrelated with standard macroeconomic variables. This, e.g., casts doubt for models such as proposed in De Long et al. (1990), relying on errors in firm valuations that are unrelated to actual cash flows, or relying on significant differences among firms' cash flows with aggregated cash flows as in Zhang (2005). Our main results are in line with Lochstoer and Tetlock (2020) and document that systematic cash flow news drives the returns of anomaly portfolios not only in the U.S., but also in the European market among EMU countries.

2.5 Anomaly cash flow and discount rate shock spillover effects

Our previous results document a common CF news component among anomaly portfolios. As shown in the previous section, this unobserved component seems to be unrelated with macroeconomic variables. However, looking at static correlation coefficients as in Section 2.4.2 may overlook any possible dynamic behavior in the time-series of CF and DR news. For this reason, we apply the time-varying parameter vector auto-regression (TVP–VAR) based extended joint connectedness framework proposed in Balcilar et al. (2021) for our CF and DR shocks, which allows for the more accurate measurement of the dynamic evolution of shocks' connectedness among anomaly portfolios. In general, the connectedness approach is based on the methods outlined in Diebold and Yilmaz (2009), Diebold and Yilmaz (2012), and Diebold and Yilmaz (2014).

To begin with, we estimate a TVP-VAR model with lag one (month):

$$y_t = B_t y_{t-1} + \epsilon_t, \quad \epsilon \sim N\left(0, \Sigma_t\right) \tag{2.9}$$

$$vec(B_t) = vec(B_{t-1}) + \upsilon_t, \quad \upsilon \sim N\left(0, R_t\right) \tag{2.10}$$

where y_t, y_{t-1}, and ϵ_t are $K \times 1$ dimensional vectors, B_t and Σ_t are $K \times K$ dimensional matrices, and R_t is a $K^2 \times K^2$ dimensional matrix, which allows all parameters (B_t) to vary over time. After transforming the TVP-VAR model to a TVP-VMA model applying the Wold representation theorem $y_t = \sum_{h=0}^{\infty} A_{h,t}\epsilon_{t-i}$, with $A_0 = I_K$, the H-step forecast error can be written as:

$$\xi_t\left(H\right) = \sum_{h=0}^{H-1} A_{h,t}\epsilon_{t+H-h}. \tag{2.11}$$

Based on Koop et al. (1996) and Pesaran and Shin (1998), the H-step ahead generalized forecast error variance decomposition (GFEVD) represents the effect a shock in series j has on series i:

$$\xi_{ij,t}^{gen}(H) = \frac{E(\xi_{i,t}^2(H)) - E\left[\xi_{i,t}(H) - E(\xi_{i,t}(H))|\epsilon_{j,t+1}, ..., \epsilon_{j,t+H}\right]^2}{E(\xi_{i,t}^2(H))}$$

$$= \frac{\sum_{h=0}^{H-1}(e_i' A_{ht}\Sigma_t e_j)^2}{(e_j'\Sigma_t e_j)\sum_{h=0}^{H-1}(e_i' A_{ht}\Sigma_t A_{ht}' e_i)} \tag{2.12}$$

$$gSOT_{ij,t} = \frac{\xi_{ij,t}^{gen}(H)}{\sum_{j=1}^{K}\xi_{ij,t}^{gen}(H)},$$

where $\xi_{ij,t}^{gen}(H)$ is the unscaled GFEVD, which is normalized to unity by dividing it by the row sum leading to the scaled GFEVD, $gSOT_{ij,t}$.

The scaled GFEVD is the fundamental on which all connectedness measure are calculated like the total directional connectedness from others to variable i and the total directional connectedness to others from a shock in variable i. This metric illustrates by how much the network influences variable i and how much variable i influences the whole network, respectively, and can be written as:

$$S_{i\leftarrow\bullet,t}^{gen,from} = \sum_{j=1,i\neq j}^{K} gSOT_{ij,t}$$
$$S_{i\rightarrow\bullet,t}^{gen,to} = \sum_{j=1,i\neq j}^{K} gSOT_{ji,t}. \tag{2.13}$$

Based upon the previous two measure, the net total directional connectedness of series i can be computed and is interpreted as the net influence of series i on the network,

$$S_{i,t}^{gen,net} = S_{i\rightarrow\bullet,t}^{gen,to} - S_{i\leftarrow\bullet,t}^{gen,from}. \tag{2.14}$$

If $S_{i,t}^{gen,net} > 0$, series i is influencing all others more than being influenced by them and thus is considered as a net transmitter of shocks indicating that series i is driving the network. In case of $S_{i,t}^{gen,net} < 0$, the series is a net receiver of shocks and thus driven by the network. At the center of the connectedness approach lies the total connectedness index (TCI) which highlights the network interconnectedness and hence average network spillover. It is calculated as the average total directional connectedness from (to) others:

$$gSOI_t = \frac{1}{K}\sum_{i=1}^{K} S_{i\leftarrow\bullet,t}^{gen,from} = \frac{1}{K}\sum_{i=1}^{K} S_{i\rightarrow\bullet,t}^{gen,to}. \tag{2.15}$$

Following Balcilar et al. (2021), we do, however, use an equivalence of $gSOT_{ij,t}$ from Equation (2.12), namely $jSOT_{ij,t}$. While the interpretation of our results does not depend on this choice, this results in more accurate estimates overcoming some shortcomings in the methods of Diebold and Yilmaz (2009), Diebold and Yilmaz (2012), and Diebold and Yilmaz (2014), as discussed in Caloia et al. (2019).

Table 2.5: Average joint connectedness of anomaly cash flow and discount rate shocks.

This table reports the average joint connectedness (in %) for five anomaly portfolios' cash flow (CF) and discount rate (DR) shocks. CF and DR shocks are estimated using the VAR approach proposed in Lochstoer and Tetlock (2020) and outlined in Section 2.3.1. The factors are denoted as in Fama and French (2015), i.e., SMB (size), HML (value), RMW (profitability), and CMA (investment). MOM denotes the momentum factor. The results for connectedness are based on the TVP-VAR(1) extended joint cennectedness approach proposed in Balcilar et al. (2021), with a 12-month-ahead generalized forecast error variance decomposition. TCI denotes the total connectedness index which highlights average network spillover effects. TO and FROM represent total directional connectedness as outlined in Section 2.5.

	SMB (CF)	HML (CF)	CMA (CF)	RMW (CF)	MOM (CF)	SMB (DR)	HML (DR)	CMA (DR)	RMW (DR)	MOM (DR)	FROM others
SMB (CF)	63.63	13.10	6.11	5.34	2.27	1.38	3.13	1.36	1.34	2.33	36.37
HML (CF)	11.00	55.24	6.96	5.18	3.79	3.18	6.16	1.55	1.89	5.04	44.76
CMA (CF)	6.53	8.74	70.07	2.22	2.81	2.01	1.32	2.76	1.78	1.75	29.93
RMW (CF)	4.51	5.58	1.20	76.59	1.57	1.76	1.87	2.92	2.10	1.90	23.41
MOM (CF)	1.31	2.78	3.67	1.67	77.38	2.81	3.22	1.18	2.68	3.28	22.62
SMB (DR)	2.16	2.89	1.49	0.94	2.27	48.56	19.03	7.76	3.17	11.72	51.44
HML (DR)	1.01	4.49	0.42	0.86	2.13	18.80	39.68	3.59	5.32	23.69	60.32
CMA (DR)	1.18	2.34	2.83	2.77	0.94	12.44	5.88	53.69	10.13	7.81	46.31
RMW (DR)	1.33	1.77	1.45	2.23	1.00	3.87	7.95	9.18	65.97	5.25	34.03
MOM (DR)	0.74	4.08	1.11	0.85	2.37	12.40	24.85	3.96	2.63	47.01	52.99
TO others	29.78	45.78	25.25	22.07	19.14	58.66	73.40	34.26	31.04	62.77	TCI
NET	−6.58	1.02	−4.68	−1.33	−3.47	7.23	13.08	−12.05	−2.99	9.78	40.22

Table 2.5 reports average results with regard to connectedness within the network of anomaly portfolio CF and DR shocks. Generally, main diagonal elements correspond to own-contribution and off-diagonal elements contribute to either "from" or "to" others. The factors in this section are denoted as in Fama and French (2015), i.e., SMB (size), HML (value), RMW (profitability), and CMA (investment), and MOM denotes the momentum factor (Carhart (1997)). Their respective CF and DR return components are obtained using the decomposition approach outlined in Section 2.3.1.

We find that the average TCI is 40.22%, which implies that, on average, 40.22% of a shock in one of anomaly CF or DR shocks spills over to all others. Generally, discount rate components among anomaly portfolio returns are the main transmitter of shocks in our network. On average, the HML DR component transmits 13.08% of shocks, followed by MOM DR (9.78%), and SMB DR (7.23%). Interestingly, this does not imply that cash flow components have to be the main receiver of shocks. While this is the case for SMB CF (6.58%), CMA CF (4.68%), and MOM CF (3.47%), the largest receiver of shocks is the discount rate component of the investment anomaly. On average, CMA DR receives 12.05% of shocks.

While the average results for connectedness are useful for a mere summary of the underlying interrelations among anomaly CF/DR shocks, they do not allow to evaluate any time-variation of connectedness. We carry on with the description of connectedness in a dynamic framework which can be seen in Figure 2.2.

Figure 2.2 shows the dynamic evolution of the TCI. Despite its peak around August 1992, reaching up to 80%, it is remarkable constant since 2005. Around the early 2000s, during times of the "Dotcom bubble" related market turmoils, we observe a moderate increase in the TCI from under 40% to nearly 50% in June 2001. Note that large TCI values typically imply high spillovers between the cash flow and discount rate shocks. In this sense, it seems to be a puzzle that the dynamic interconnectedness is not remarkable influenced by economic events like the global financial crisis around 2008 or the Covid-19 related stock market turmoils around 2020. Again, we already observed that CF and DR shocks are only weakly linked with shocks to macroeconomic variables, and our dynamic results suggest that related spillovers are quite independent from notable economic events, too. Looking at the static results for interconnectedness, however, this might be the case because DR anomaly shocks tend to spill over among anomalies, but are absorbed within the category of DR shocks. Further, our results in Section 2.4.1 emphasized that CF shocks tend to dominate the returns of anomaly portfolios, but DR shocks seem to play

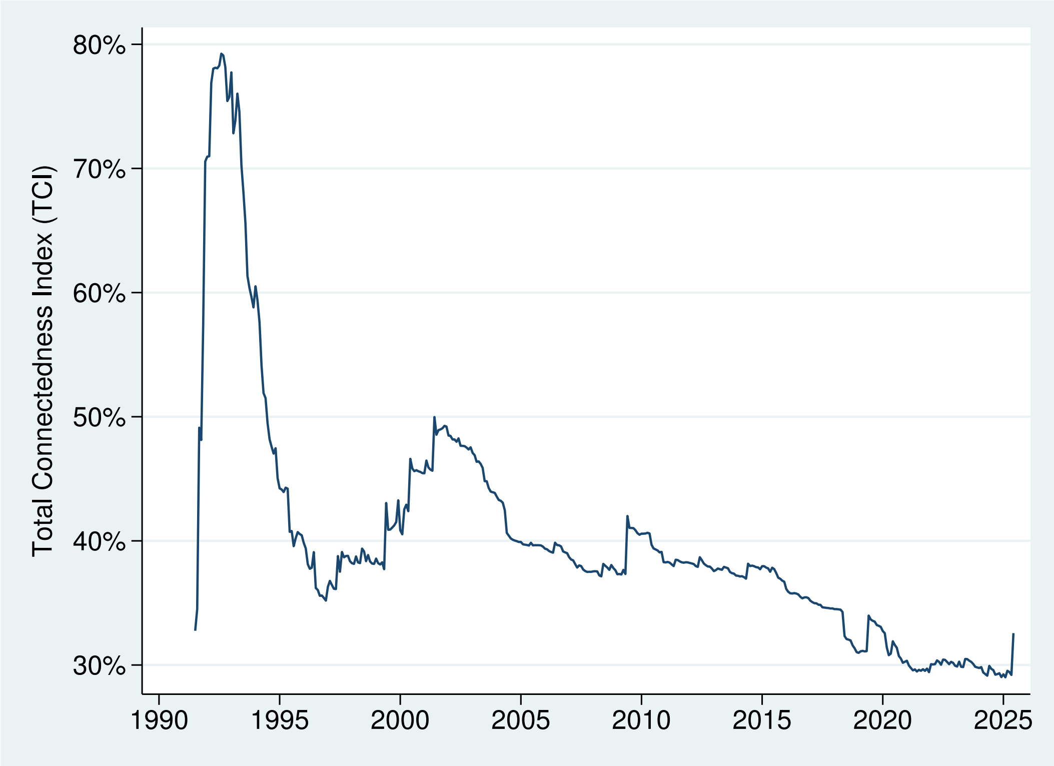

Fig. 2.2. This figure shows the total connectedness index (TCI) based on a TVP-VAR model proposed in Balcilar et al. (2021) with lag length of order one (month) and a 12-month-ahead generalized forecast error variance decomposition. We consider cash flow and discount rate shocks for five anomaly portfolios according to Fama and French (2015) (size, value, investment, and profitability) and Carhart (1997) (momentum) in the model. Cash flow and discount rate shocks are estimated using the decomposition approach proposed in Lochstoer and Tetlock (2020) and outlined in Section 2.3.1. The sample period is June 1991 to June 2025 and the sample includes firms located in the European Monetary Union.

an important role in carrying shocks from one anomaly to the others.

2.6 Conclusion

In this paper, we examine the returns of anomaly portfolios and decompose their returns into cash flow and discount rate components. Relying on the decomposition approach proposed in Campbell and Shiller (1988a), Campbell (1991), and more recent, Lochstoer and Tetlock (2020), we dissect the returns for the size, value, investment, profitability, and momentum anomaly in the European Monetary Union from June 1991 until June 2025. We provide strong supportive out-of-sample-evidence on the U.S. findings by demonstrating that cash flow shock variation clearly dominates the returns of not only the anomaly portfolios, but also their mean-variance efficient combination. Anomaly cash flow shocks are occasionally correlated with shocks to aggregate macroeconomic variables (mainly consumer confidence, default spread, and M2 money supply), while anomaly discount rate shocks are largely uncorrelated with them. Market cash flow shocks are negatively correlated with shocks to the default spread, which is a plausible measure of risk aversion. Similarly, market discount rate shocks are negatively correlated with consumer confidence and export shocks, implying that market discount rates increase in recessions. While this is as expected, cash flow and discount rate shocks of the mean-variance efficient portfolio comprising our five anomaly portfolios are largely uncorrelated with macroeconomic shocks.

Applying the TVP-VAR connectedness framework proposed in Balcilar et al. (2021) to evaluate spillover effects, we find that, on average, 40.22% of a shock in one of anomaly CF or DR shocks spills over to all others. The largest shocks are obtained within discount rates, but their spillovers to the CF shock category seems to be modest. Despite the dominance of CF shocks among anomaly portfolio returns, DR shocks seem to play a more important role in carrying shocks from one anomaly to the others. These results provide guidance for asset pricing models to account for the empirical facts presented in this study, mainly (i) that cash flow shocks show a common component among anomaly portfolios, (ii), cash flow shocks show little relation to the business cycle, and (iii), discount rates are the main channel for transmitting spillover effects to other anomalies.

Chapter 3

Pricing and Mispricing of Accounting Fundamentals: Global Evidence

This research is the result of a single-author project. The paper has been published as:

Köstlmeier S. (2024), Pricing and mispricing of accounting fundamentals: Global evidence, *The Quarterly Review of Economics and Finance* (94), p. 71-87.

The journal ranking is *B* according to the VHB Publication Media Rating 2024. The paper was presented at the poster session of the 29th Annual Meeting of the German Finance Association (DGF) 2023 in Hohenheim.

Abstract

This paper extends the fundamentals-based valuation model in Nichols et al. (2017) to global, developed equity markets. The model is able to explain, on average, 81% of the cross-sectional share price variation among global stocks. To be applicable among international markets, actual cash-flow streams instead of clean surplus accounting figures are used to reflect the different importance of dividends and share repurchases around the world. Firms identified as undervalued outperform overvalued firms by 0.62% p.m. after controlling for size, book-to-market, operating profitability, investment, and momentum. This premium is further not explained by lottery-like stock preferences (MAX, idiosyncratic volatility, skewness), mispricing related variables (FSCORE, ΔXFIN), or stock issuances. In support of a mispricing related explanation, we detect a significant post publication return decline in the easily exploitable long portfolio leg comprising undervalued stocks. Together with our analysis on investor sentiment, portfolio transitions, and arbitrage asymmetry, we provide evidence that deviations of the share price from the model's estimated value indicate actual mispricing and according returns are unlikely to be a compensation for risk exposure.

Keywords: Empirical asset pricing · Mispricing · Fundamental analysis · International markets

JEL classification: G11 · G12 · G15.

3.1 Introduction

In their seminal work on a cross-sectional fundamentals-based firm valuation model, Nichols et al. (2017) find a strong relation between the absolute share price and accounting fundamentals. They find that at least 63% of the cross-sectional share price variation of U.S. firms (1975-2011) can be traced back to variations in the book value of equity, income, dividends, and growth in operating income. Walkshäusl (2021) presents strong out-of-sample evidence for firms (1990-2017) located in European Monetary Union (EMU) countries by explaining 69% of their share price variation. The valuation model directly addresses concerns stated by J. Cochrane in his 2011 AFA presidential address that the field of asset *pricing* has in fact evolved into asset *expected returns* (Cochrane (2011)).

An important caveat is that the aforementioned studies exclude firms with fiscal year-ends other than the end of December. Even though December 31 is the most popular fiscal year-end in the U.S., this a priori excludes 30% of all firms from the sample. Within EMU countries, we find that 86% of firms choose December 31. However, only 43% of firms from the United Kingdom use that specific date, though extending the sample to the whole region of Europe would already exclude more than half of according firm observations. Fiscal year-ends generally differ a lot around the world: Japanese firms prefer the end of March (70%) while only 9% use December 31. Within the region of Asia-Pacific, 40% use June 30 and 38% December 31. Including these firms in our analysis seems to be crucial, as any findings in empirical research could be the result of data snooping as shown by Lo and MacKinlay (1990) or be sample specific in the meaning of Harvey and Liu (2016).

Previous studies are further based on clean surplus assumptions, i.e., the change in book value of equity between two periods is assumed to equal the difference between earnings and dividends. The latter are calculated as a residual figure and are considered to be the single mean of distribution to investors within the valuation model. Ohlson (2005) emphasizes that this approach is unbiased only if the expected equity transactions are neutral or irrelevant to prospective new shareholders. This seems rather unlikely given the fact that share repurchases have become more important for distributing wealth to shareholders than dividends in recent times (Denis and Osobov (2008); Fama and French (2008a)). While this is true for the U.S., following accounting practices emphasizing fair-market valuations and a capital market perspective, dividends still remain an important source of non-capital gains in countries following accounting systems that reflect a traditional banking orientation (Booth et al. (2001) and Hung (2000)). In a parsimonious model as presented in Nichols et al. (2017), any deviance of clean surplus assumptions could have a potential impact on results, especially when analyzing a global sample of firms located in countries with different accounting principles as in this study.[1]

In this study, we propose an extended version of the fundamentals-based firm valuation model that separately accounts for dividends and share repurchases and we demonstrate how to include the whole spectrum of fiscal year-ends into an aggregated, quarterly updated model.

Our results are easily summarized. Accounting fundamentals are able to explain on average 81% of share price variation among a global sample of 15,617 non-U.S. firms. We document a strong positive relation of actual cash-dividends paid and a strong negative relation of share repurchases with according share prices. Following Nichols et al. (2017), we examine the ability of deviations of actual share prices from their model's estimates, denoted as value residual (*VRES*), to predict future stock returns. Sorting stocks into global quintile portfolios based on *VRES* shows a size-adjusted outperformance of low *VRES* (i.e., undervalued firms) of 0.56% per month. These undervalued firms, on average, tend to be smaller (avg. market capitalization equals $633.02m), have a higher book-to-market ratio (1.21), a weaker profitability (0.72) and lower momentum (0.08), compared with overvalued firms. Our cross-sectional regression analysis documents that the return difference between low and high-*VRES* firms remains economically and statistically significant and amounts to 0.62% per month after controlling for firm size, book-to-market ratio, operating profitability, investment, and momentum. Overall, our results provide convincing evidence for the main findings in Nichols et al. (2017) in global, developed capital markets around the world.

The remainder of this study is organized as follows. Section 3.2 describes the data and variables used in this study. Section 3.3 discusses how the fundamentals-based valuation model proposed in Nichols et al. (2017) can be extended to be applicable to a global sample of common stocks. Section 3.4 and 3.5 contains our main analysis and further robustness tests. Finally, Section 3.6 concludes the paper.

3.2 Data and summary statistics

We use our extension of the fundamentals-based valuation model to explore the cross-sectional return variation among 21 global, developed non-U.S. equity markets according

[1] This is the reason why the original study finds an insignificant dividend-price relation in the U.S. while Walkshäusl (2021) who analyzes firms located in the European Monetary Union reports a significantly positive coefficient estimate for dividends.

to the MSCI classification. Our sample comprises on average 4,542 firms each year from June 1990 to June 2021.[2] We obtain monthly financial data on common stocks from Refinitiv Datastream and accounting information from Worldscope. All data is denominated in currency U.S.-$. To ensure that our results are not driven by very small and illiquid firms, we follow Ang et al. (2009) and eliminate the 5% of firms with lowest market capitalization measured at the end of June in each country from our sample. To avoid a look-ahead bias, accounting information of the latest fiscal year ending in the previous calendar year is matched with financial data from July of the current year to June of the subsequent year. We follow Ince and Porter (2006), Griffin et al. (2010) and Schmidt et al. (2017) and apply the suggested static and dynamic screens to ensure a high data quality and to exclude non-common equity securities like ADRs, Funds or REITS. We delete monthly returns greater than 990% or abnormal high returns followed by strong reversal as in Ince and Porter (2006).[3] According to Fama and French (1992) we exclude financial firms indicated by an SIC code between 6,000 and 6,999 or an ICB industry number of 30 or 35. If the SIC code or ICB industry number is missing, we disregard the firm.

All variables used in this study are winsorized at the 1%/99% levels and defined as follows. As in Fama and French (2015), firm size ($SIZE$) is the market value of equity at the end of June of each year in million dollars. Book-to-market ratio ($BMRATIO$) is the ratio of book equity to market equity for the fiscal year ending in the previous calendar year. Investment (INV) is the annual growth of total assets and operating profitability (OP) is revenues minus cost of goods sold and interest expense, all divided by book equity.[4] The cumulative prior twelve-month stock return (Jegadeesh and Titman (1993)), skipping the most recent month, is denoted as momentum (MOM). Short-term reversal (REV) is the stock return over the previous month (Jegadeesh (1990)). As in Fama and French (2008b), accrual (ACC) is the annual change in operating working capital per split-adjusted share divided by book equity per split-adjusted share and net stock issue (NSI) is the change in the natural log of split-adjusted shares outstanding. $\Delta XFIN$ is a financing-based measure of misvaluation proposed in Bradshaw et al. (2006). It is the sum of net equity financing and net debt financing all divided by lagged total assets. The fundamental strength of firms ($FSCORE$) is measured as the sum of nine binary

[2]The sample period for Israel starts in June 2000 because of data availability constraints.

[3]If r_{t-1} or r_t is greater than 300% per month and $(1 + r_{t-1}) \cdot (1 + r_t) - 1 < 0.5$, we eliminate both return observations.

[4]As in Walkshäusl (2021), we do not include selling, general, and administrative expenses because this data item is not broadly available among non-U.S. firms.

variables equal to one, if the according conditions following Piotroski (2000) are met. MAX is the maximum daily stock return over the previous month (Bali et al. (2011)) and $SKEW$ is computed as total skewness of daily stock returns over the previous month. Idiosyncratic volatility ($IVOL$) is the annualized idiosyncratic volatility relative to the Fama-French model using daily stock returns over the previous month (Ang et al. (2006) and Ang et al. (2009)).[5] Reflecting modern developments in empirical asset pricing, we apply the Fama-French five-factor model at the regional level to calculate $IVOL$ (Fama and French (2015); Hollstein (2022)).[6]

The fundamentals-based valuation model presented in Nichols et al. (2017) and our extended version rely on further variables defined as follows. $PRICE$ denotes the absolute share price and BV is book value of equity per shares outstanding. IB is income before extraordinary items divided by shares outstanding. NEG is a binary indicator variable that is equal to one if IB is negative and zero otherwise. According to Nichols et al. (2017), DIV^{cs} is dividends per shares outstanding based on clean surplus accounting. It is measured as beginning book value of equity less ending book value of equity plus net income before extraordinary items and preferred and common dividends, all divided by shares outstanding. In contrast, DIV^{act} is total cash dividends paid divided by shares outstanding. $OIGR$ denotes the growth in operating income per share and is calculated as the annual change of operating income divided by shares outstanding. BUY is net repurchase financing divided by shares outstanding for share repurchasing firms and zero otherwise. Share repurchasing firms are identified by having a net repurchase financing greater than 0.1% of the market value of equity as measured at the end of the previous fiscal year (see Dittmar (2000), Grullon and Roni (2002), and Jacob and Jacob (2013)). Net repurchase financing is the purchase of common and preferred stock minus any reduction in the book value of preferred stock (if available and zero otherwise) that occurs between year $t-1$ and year t.

[5] We require a minimum of 16 daily stock returns per month for all variables using daily stock returns.

[6] We obtain regional (Europe, Asia-Pacific, Japan) factor returns from Kenneth French's website and are grateful for making these data available.

Table 3.1: Descriptive statistics, June 1990 - June 2021.

This table shows summary statistics for the countries included in our global sample and the variables used in this study. Panel A reports the average number of firms in each country and the country's average percentage weight in terms of total market equity. Panel B reports time-series averages of the cross-sectional mean, standard deviation, 25th percentile, median and 75th percentile of the variables for the sample period from June 1990 (Israel: June 2000) to June 2021. *SIZE* is the market value of equity at the end of June of each year in million U.S. dollars. *BMRATIO* is the book value of equity at the fiscal year-end divided by the market value of equity at the end of December. *INV* is the annual change in total assets divided by prior year's total assets. *OP* is revenues minus cost of goods sold and interest expense, all divided by book value of equity. *MOM* is the cumulative prior twelve-month stock return skipping the most recent month. *REV* is the stock return over the previous month. *ACC* is the annual change in operating working capital per split-adjusted share divided by book equity per split-adjusted share. *NSI* is the change in the natural log of split-adjusted shares outstanding. $\Delta XFIN$ is defined as the sum of net equity financing and net debt financing divided by lagged total assets as proposed in Bradshaw et al. (2006). *FSCORE* is a composite measure of the firm's fundamental strength as in Piotroski (2000). *IVOL* is the annualized idiosyncratic volatility relative to the Fama–French five-factor model using daily stock returns over the previous month. *SKEW* is total skewness using daily stock returns over the previous month. *MAX* is the maximum daily stock return over the previous month. *PRICE* is the share price denoted in U.S.-\$. *BV* is the book value of equity per shares outstanding. *IB* is income before extraordinary items divided by shares outstanding. *NEG* is a binary indicator variable that is equal to one if *IB* is negative and zero otherwise. *OIGR* is the annual change of operating income divided by shares outstanding. *BUY* is net repurchase financing of common and preferred stocks divided by shares outstanding for share repurchasing firms and zero otherwise as defined in Dittmar (2000). DIV^{act} is total cash dividends paid divided by shares outstanding.

Panel A: Developed sample countries

Country	Firms	Weight	Country	Firms	Weight
Australia	427	3.86	Japan	1,306	28.07
Austria	23	0.33	Netherlands	73	3.01
Belgium	37	1.11	New Zealand	39	0.29
Denmark	68	1.05	Norway	71	1.05
Finland	66	1.32	Portugal	26	0.35
France	262	10.08	Singapore	224	1.28
Germany	214	7.94	Spain	57	2.63
Hong Kong	434	5.62	Sweden	139	2.27
Ireland	21	0.44	Switzerland	103	5.78
Israel	83	0.56	United Kingdom	752	20.29
Italy	117	2.86			

Panel B: Variables

Variable	Mean	St. Dev.	25th	Median	75th
$SIZE$	1,895.00	8,143.73	55.82	194.52	809.86
$BMRATIO$	0.94	0.82	0.40	0.71	1.20
INV	0.13	0.37	−0.04	0.05	0.17
OP	0.75	0.85	0.28	0.53	0.94
MOM	0.11	0.45	−0.16	0.04	0.28
REV	0.01	0.11	−0.05	0.00	0.06
ACC	−0.05	0.47	−0.07	0.00	0.07
NSI	0.08	0.31	0.00	0.00	0.03
$\Delta XFIN$	0.05	0.24	−0.04	−0.01	0.04
$FSCORE$	5.65	1.65	4.56	5.73	6.81
$IVOL$	0.30	0.22	0.16	0.24	0.35
$SKEW$	0.18	1.03	−0.34	0.17	0.71
MAX	0.06	0.05	0.03	0.04	0.07
$PRICE$	20.53	55.93	1.14	5.01	15.30
BV	13.92	38.85	0.70	3.07	10.50
IB	0.78	3.59	0.02	0.18	0.74
NEG	0.21	0.39	0.00	0.00	0.26
$OIGR$	0.05	2.44	−0.10	0.01	0.17
BUY	0.04	0.14	0.00	0.00	0.00
DIV^{act}	0.37	1.06	0.01	0.07	0.25

Distributional statistics for our global sample are presented in Table 3.1. On average, a typical firm has a market capitalization of $1.9b, a book-to-market ratio of 0.94 and total asset grow 13% per year. The average share price of $20.53 is composed of $13.92 book value of equity which implies that the remaining $6.61, resp. 32% of the average share price, are attributable to expected future residual income.

3.3 Extending the fundamentals-based valuation model

The fundamentals-based valuation model proposed in Nichols et al. (2017) considers only U.S. firms who's fiscal year-end is December 31. Share prices at the end of subsequent March are regressed on accounting fundamentals applying annual firm-level cross-sectional regressions:

$$\begin{aligned} PRICE_{i,t} = &\sum_{j=1}^{9} a_{j,t} + \gamma_{1,t} BV_{i,t} + \gamma_{2,t} IB_{i,t} + \gamma_{3,t} NEG_{i,t} + \gamma_{4,t} \left(NEG \times IB\right)_{i,t} \\ &+ \gamma_{5,t} DIV^{cs}_{i,t} + \gamma_{6,t} OIGR_{i,t} + \epsilon_{i,t}. \end{aligned} \tag{3.1}$$

$a_{j,t}$ is a binary indicator for nine distinct industry groups and takes the value of one if a firm belongs to the given industry and zero otherwise. Fitted values are denoted as $VALUE$ and indicate fundamentals related share prices. As in Nichols et al. (2017) and Walkshäusl (2021), negative $VALUE$ estimates are not meaningful and typically excluded from the sample. Deviations of the actual share price from the model's estimate are referred to as value residual ($VRES$) and negative values indicate a firm's undervaluation relative to its firm fundamentals:

$$VRES_{i,t} = \frac{PRICE_{i,t} - VALUE_{i,t}}{PRICE_{i,t}}. \tag{3.2}$$

Although December 31 is the most popular fiscal year-end in the U.S. and EMU, extending the original model as stated in Eq. (3.1) to other capital markets around the world induces two main problems:

First, Japanese firms prefer the end of March as fiscal year-end (70%) while only 9% use December 31. Similarly, only 38% of firms located in Asia-Pacific chose December 31. In Europe, most firms are located in the UK, and only 43% of them chose December 31. Extending the sample from EMU countries to Europe would a priori exclude more than half of firm observations located in the UK. However, including these firms in our analysis seems to be crucial, as any findings in empirical research could be the result of data snooping as shown by Lo and MacKinlay (1990) or be sample specific in the meaning of Harvey and Liu (2016).

Second, we demonstrate that the original model has to be adjusted to be meaningful for a global sample of firms relying on different accounting principles. To be specific, what is labeled as *dividends* in Nichols et al. (2017) (here: DIV^{cs}) is actually an indirectly measured, residual figure based on clean-surplus accounting assumptions. In consequence, DIV^{cs} neglects the growing influence of share repurchases in recent times, especially for firms located in Anglo-Saxon capital market systems. In the manner of Ohlson (2005), DIV^{cs} is unbiased only if the expected equity transactions are neutral or irrelevant to prospective new shareholders, which is typically not the case for share repurchasing firms.

To demonstrate our point, we consider a European sub-sample in the following section for two reasons: First, our study is related with Walkshäusl (2021) who only considers EMU located firms until 2017, so all according findings are directly comparable and not sample specific.[7] Second, the region of Europe can easily be parted into a sub-sample of EMU countries, following accounting systems that reflect a banking orientation, and the United Kingdom, following accounting practices emphasizing fair-market valuations and a capital market perspective (see Booth et al. (2001); Hung (2000)).

3.3.1 Dividends, share repurchases, and the clean-surplus accounting assumption

Looking at the European sub-sample, we observe on average 2,871 firms per year and most of them are located in the United Kingdom, France, or Germany. Together, these countries contribute to a total of 62% of aggregated market capitalization in Europe. Firms located in non-EMU countries distribute on average €0.07 per share to investors as measured by DIV^{cs} which is based on clean-surplus accounting assumptions. In contrast, EMU firms distribute €0.41 and the according difference of €0.34 is highly significant with a Newey and West (1987) robust t-statistic of 2.64. However, there is no significant difference with regards to actual cash dividends paid (DIV^{cs}). The average share repurchase volume for non-EMU firms amounts to €0.04 per share, resp. €0.05 for EMU firms. Given the fact that the average share price for European, non-EMU firms is much smaller (€23.54) compared with EMU firms (€39.12), share repurchases are economically more important

[7] All data in Section 3.3.1 are denoted in currency euro for that reason.

in the UK and related countries.[8]

Dividends are the most important transfer of wealth to shareholders in Europe compared to the U.S., where share repurchases are more favorable than dividends in recent years (see e.g., Denis and Osobov (2008); Fama and French (2008a)). We observe an aggregated buyback volume of €833.08b in Europe (excl. UK) in the thirty-one years from 1990 to 2021, similar to Fried and Wang (2021). Of that volume, €475.89b were repurchased within the most recent decade. Nevertheless, actual cash dividends paid comprised an aggregate value of €3.42t and from that, €2.12t were paid during the last ten years. In the UK, the aggregated buyback (dividend) volume is €500.08b (€1.64t) in total and €239.65b (€887.49b) between 2011 and 2021. In summary, share repurchases in Europe (incl. UK) account for an average of 19% of all payouts in the last decade. In the first half of our sample period (1990-2005), only 6% of all European firms repurchased own shares, compared with 19% for the period 2006-2021. This relative importance highlights the necessity to separately account for share buybacks and dividends in cross-sectional regressions. Because a firm's choice between paying out cash flows in form of dividends or stock repurchases depends on different economic motives (e.g., financial flexibility, see Jagannathan et al. (2000)), we should dissect DIV^{cs}.

Spec. (1) in Table 3.2 replicates the original valuation model (see Eq. (3.1)) but dissects the effect of DIV^{cs} for EMU and non-EMU located firms.[9] We observe a significant positive relation between DIV^{cs} and stock prices for EMU located firms as previously documented in Walkshäusl (2021). However, non-EMU firms seem not to price DIV^{cs} because the net effect of 0.11 (formally, $1.83 + (-1.73)$) is statistically and economically insignificant. In comparison, Spec. (3) and (4) use actual cash dividends paid and share repurchases as separate variables. A €1.00 increase in DIV^{act} is associated with a higher stock price of €5.96. Share repurchases in Europe are unrelated with future stock prices on average, but for Anglo-Saxon capital markets as the UK, we find that an additional €1.00 repurchase volume per share decreases the average share price by €41.14. To put that into perspective, the average (European) firm has a share repurchase volume of €0.04 per share with a cross-sectional standard deviation of €0.20, so a one standard deviation increase of the average buyback volume decreases the share price by €8.23. In conclusion, share repurchases are an economically important driver of stock prices in the

[8] The sample mean of €0.04 takes into account non-repurchasing firms that have a value of zero for BUY and prevail in our sample. BUY for share repurchasing firms is on average €0.24 which implies that the annual buyback volume equals 0.79% of a firm's market capitalization.

[9] We already apply our extended model approach described in Section 3.3.2 which allows us to include all firms from our sample no matter their fiscal year-end.

Table 3.2: Cross-sectional regressions (Europe) for the valuation model in Nichols et al. (2017).

This table shows time-series averages for the coefficients from firm-level cross-sectional regressions of the share price on accounting fundamentals. BV is the book value of equity, IB is income before extraordinary items, DIV^{cs} is book value of equity in year y_{t-1} minus book value of equity in year y_t plus net income in year y_t, DIV^{act} is cash dividends paid, and $OIGR$ is the annual change of operating income. All of the aforementioned variables are divided by shares outstanding. NEG is a binary indicator variable that is equal to one if IB is negative and zero otherwise. BUY is the net share repurchase volume per shares outstanding. If a firm has a net repurchase volume less than 0.1% of the market value of equity at the previous fiscal year-end or does not repurchase stock at all, BUY is set to zero. *nonEMU* is a binary variable set to one if the according company is located in a European country not part of the European Monetary Union and zero otherwise. All specifications include industry dummies based on Industry Classification Benchmark (ICB) categories, country dummies, and dummie indicators for the month of the fiscal year-end. We conduct quarterly cross-sectional regressions using firms having their fiscal year-end in the previous three months. To avoid a look-ahead bias, our dependent variable (share price) is measured using a lag of three months following the end of the calendar quarter. The R^2 values are adjusted for degrees of freedom. N denotes the average number of firms. *Region* indicates the countries included in the sample. Newey and West (1987) corrected t-statistics are given in parenthesis and */**/*** indicate significance at the 10%/5%/1% level. The sample period is from June 1990 to June 2021.

Specification	(1)	(2)	(3)	(4)	(5)
BV	0.68***	0.82***	0.65***	0.90***	0.59***
	(14.29)	(9.86)	(11.94)	(12.51)	(11.18)
IB	8.89***	9.12***	7.78***	6.96***	7.98***
	(14.68)	(11.74)	(11.79)	(8.12)	(12.77)
NEG	−1.62***	−0.02	−0.96**	−0.18	−0.98**
	(−3.61)	(−0.08)	(−2.06)	(−0.67)	(−2.15)
$NEG \times IB$	−14.54***	−13.27***	−13.42***	−12.19***	−13.30***
	(−13.94)	(−9.49)	(−13.33)	(−6.30)	(−13.36)
DIV^{cs}	1.83***	−0.36			0.50
	(2.59)	(−1.37)			(0.50)
$DIV^{cs} \times nonEMU$	−1.72**				−1.03
	(−2.13)				(−0.97)
DIV^{act}			5.96***	5.71***	6.96***
			(3.77)	(4.39)	(4.46)
$DIV^{act} \times nonEMU$			−1.82		−1.89
			(−1.05)		(−1.14)
BUY			32.31	−10.05**	27.31
			(1.05)	(−2.41)	(0.97)
$BUY \times nonEMU$			−73.45**		−65.04**
			(−2.24)		(−2.18)
$OIGR$	−0.10	1.54***	0.78**	2.02***	0.27
	(−0.29)	(3.10)	(2.13)	(4.18)	(0.76)
$Adj.R^2$	0.87	0.83	0.88	0.83	0.89
N	731	228	722	231	713
Region	Europe	UK	Europe	UK	Europe

fundamentals-based valuation model, however, dismissed in its original version.

Finally, Spec. (5) supports our findings that DIV^{cs} is not suitable for analyzing a global sample of firms with different fundamental accounting practices, although the average cross-sectional Spearman correlation between DIV^{cs} and DIV^{act} is 0.50. While the average coefficient for DIV^{cs} is insignificant, DIV^{act} and BUY have strong explanatory power for subsequent share prices in European equity markets. Notably, we observe contrary signs for both coefficient estimates which highlights the need to replace DIV^{cs} by them. Why do both payout channels exhibit opposite economic relationships despite both forms represent ways to distribute capital to shareholders? Based on the existing literature, we suggest a firm life cycle dependent payout policy with two different economic channels as follows:[10] Jacob and Jacob (2013) show among 25 analyzed countries that relative firm size increases with the likelihood that a firm pays dividends. Regular dividend payments further establish trust among investors (see e.g., DeAngelo and DeAngelo (2006), Guiso et al. (2008), Ham et al. (2020), and Kapons et al. (2023)), and signal positive information about future cash flows (see e.g., Brickley (1983), Nissim and Ziv (2001), and Michaely et al. (2021)) - especially for small firms. Taken together, we expect a positive dividend-price relation as documented in Walkshäusl (2021) for firms in early, non-matured stages of their life cycle. In contrast, mature, highly profitable firms that already pay dividends and whose equity capitalization consists mostly of retained earnings are more likely to repurchase shares (DeAngelo et al. (2006)). These mature firms often have low future growth opportunities and high levels of cash because of limited investment opportunities.[11] In consequence, higher repurchases imply lower reinvestment in the business, lower future growth, and hence lower share prices.

Our data provides supporting evidence for these firm life cycle dependent payout mechanisms: The average market capitalization for non-repurchasing, non-dividend paying firms is only \$425m but increases to \$2.13b for dividend paying firms and to a very high \$4.41b for repurchasing, dividend paying firms. Non-dividend paying, repurchasing firms have higher cash holdings (18% of total assets) than dividend paying firms (15%), are less profitable (0.70 vs. 0.76), are fundamentally weaker (5.64 vs. 5.87) as indicated by the F-SCORE measure proposed by Piotroski (2000), and have a higher book-to-market ratio (1.02 vs. 0.94). In line with our argument, dividend paying firms tend to become share

[10] Firm life cycle is a key determinant for equity valuation and a strong predictor of futures stock returns as shown in Konstantinidi (2022) and the literature therein.

[11] Both characteristics are main motives for a firm to repurchase stocks and the literature documents a positive (negative) relation between repurchase activities and measures of excess cash (growth opportunities), see e.g., Easterbrook (1984), Jensen (1986), Dittmar (2000), or Grullon and Roni (2004).

repurchasing firms over time. Less than a fifth of dividend paying firms simultaneously repurchase own shares. On the other hand, 89% of share repurchasing firms pay cash dividends to investors. A probit regression of an indicator variable being one for share repurchasing, dividend paying firms in fiscal year $t+1$ and zero otherwise on an indicator variable being one for non-repurchasing, dividend paying firms in fiscal year t reveals, that already being a dividend paying firm increases the probability of initiating share repurchases in the following fiscal year by 15%.[12]

3.3.2 The fundamentals-based valuation model in global equity markets

We adjust the valuation model proposed in Nichols et al. (2017) based on our insights from the previous section that one should account for the different importance of dividends and share repurchases among global capital markets. For a more in depth analysis, we further apply quarterly instead of annual cross-sectional regressions which allows us to include all firms in our global sample no matter their fiscal year-end:[13]

$$\begin{aligned} PRICE_{i,t} = & \sum_{j=1}^{9} a_{j,t} + \gamma_{1,t} BV_{i,t} + \gamma_{2,t} IB_{i,t} + \gamma_{3,t} NEG_{i,t} + \gamma_{4,t} \left(NEG \times IB\right)_{i,t} \\ & + \gamma_{5,t} DIV_{i,t}^{act} + \gamma_{6,t} BUY_{i,t} + \gamma_{7,t} OIGR_{i,t} + \text{Country Dummies}_{i,t} \\ & + \text{Fiscal Month Dummies}_{i,t} + \epsilon_{i,t}. \end{aligned} \tag{3.3}$$

At the end of each calendar quarter, we consider all firms having their fiscal year-end within the previous three months. To avoid a look-ahead bias, our dependent variable (share price) is measured using a lag of three months following the end of a specific calendar quarter. For instance, our first quarterly regression only considers firms having fiscal year-ends in January, February, or March and matches according accounting data with share prices as of the end of June.

[12]The estimated coefficient of 1.13 is highly significant with a t-statistic of 24.31 using robust standard errors clustered by firm and fiscal year.

[13]This approach, however, still relies on annual accounting data, because quarterly data is generally not available for non-U.S. firms prior to 2001.

Table 3.3: Cross-sectional regressions (global) for the valuation model in Nichols et al. (2017).

This table shows time-series averages for the coefficients from firm-level cross-sectional regressions of the share price on accounting fundamentals. BV is the book value of equity, *IB* is income before extraordinary items, and *DIV*^{act} is cash dividends paid. *OIGR* is the annual change of operating income. All of the aforementioned variables are divided by shares outstanding. *NEG* is a binary indicator variable that is equal to one if *IB* is negative and zero otherwise. BUY is the net share repurchase volume per shares outstanding. If a firm has a net repurchase volume less than 0.1% of the market value of equity at the previous fiscal year-end or does not repurchase stock at all, BUY is set to zero. *nonEMU* is a binary variable set to one if the according company is located in a European country not part of the European Monetary Union and zero otherwise. All specifications include industry dummies based on Industry Classification Benchmark (ICB) categories, country dummies, and dummie indicators for the month of the fiscal year-end. We conduct quarterly cross-sectional regressions using firms having their fiscal year-end in the previous three months. To avoid a look-ahead bias, our dependent variable (share price) is measured using a lag of three months following the end of the calendar quarter. The R^2 values are adjusted for degrees of freedom. N denotes the average number of firms. *Region* indicates the countries included in the sample. Newey and West (1987) corrected t-statistics are given in parenthesis and */**/*** indicate significance at the 10%/5%/1% level. All variables are denominated in currency U.S.-$. The sample period is from June 1990 to June 2021.

Specification	(1)	(2)	(3)	(4)	(5)
BV	0.66^{***}	0.72^{***}	0.66^{***}	0.83^{***}	0.77^{***}
	(12.78)	(9.35)	(12.05)	(18.95)	(17.89)
IB	7.58^{***}	9.29^{***}	5.08^{***}	8.34^{***}	6.56^{***}
	(11.88)	(11.00)	(5.35)	(21.29)	(16.06)
NEG	-1.27^{**}	-0.66	-0.05	-1.03^{***}	-1.12^{***}
	(-2.31)	(-0.77)	(-0.26)	(-2.62)	(-2.93)
$NEG \times IB$	-13.05^{***}	-18.39^{***}	-4.56^{*}	-15.42^{***}	-14.03^{***}
	(-13.37)	(-5.16)	(-1.72)	(-9.44)	(-9.24)
DIV^{cs}				0.06	
				(0.28)	
DIV^{act}	5.79^{***}	14.16^{***}	6.49^{***}		8.67^{***}
	(3.76)	(4.93)	(3.23)		(8.55)
$DIV^{act} \times nonEMU$	-2.00				
	(-1.15)				
BUY	37.57	19.56^{*}	-0.55		-74.72^{**}
	(1.15)	(1.95)	(-0.05)		(-2.38)
$BUY \times nonEMU$	-81.66^{**}				
	(-2.30)				
$OIGR$	0.91^{**}	1.72^{***}	1.03	0.03	0.78^{***}
	(2.55)	(3.19)	(1.46)	(0.12)	(3.62)
$Adj.R^2$	0.87	0.84	0.83	0.80	0.81
N	722	645	427	1,727	1,720
Region	Europe	Japan	Asia Pacific	All	All

Table 3.3 reports coefficient estimates for each region Europe, Japan, and Asia-Pacific, as well as for the global sample. The majority of accounting fundamentals used in Nichols

et al. (2017) have strong significant explanatory power for the absolute share price in each region. Spec. (4) and (5) supplements our findings from the previous section that using DIV^{cs} as a proxy for dividends is not appropriate when analyzing an integrated, global sample. We observe that an additional dollar of cash dividends paid is associated with a higher share price of \$8.67. If a firm would increase average share repurchases by one standard deviation, the share price would decrease by \$10.46 ($0.14 \times (-74.72)$). In contrast to findings for the U.S., we find a positive relationship between operating income growth and share prices in Europe and Japan. A \$1.00 increase in operating income growth contributes \$0.78 to the share price at the global level. Overall, the extended valuation model represents a vital workhorse for capital markets around the world as indicated by the high value of 0.81 for the adj. R^2 which varies between 0.58 (Q3 2004) and 0.95 (Q2 1990) over time.

3.4 The fundamentals-based valuation model and subsequent stock returns

We present our results for the relation between the value residual *VRES* and future stock returns in two subsections. First, we use extensive portfolio sorts and cross-sectional regressions in the manner of Fama and MacBeth (1973). Second, we apply time-series regressions to examine if the according return premium earned by undervalued firms is actually useful in pricing the cross-section of stock returns.

3.4.1 Portfolio analysis

To examine how global stock returns vary with different levels of *VRES*, we begin our analysis at the portfolio level. We form quintile portfolios at the end of each calendar quarter by allocating firms in ascending order to five groups based on *VRES*.[14] Monthly portfolio returns are calculated for the subsequent three months, and the portfolios are rebalanced each quarter. We calculate equal-weighted, value-weighted, and size-adjusted portfolio returns. For the size-adjustment, the monthly raw return is measured net of the (equal-weighted) return on its matching size quintile portfolio.

Table 3.4 presents monthly portfolio returns along with average firm characteristics for quintile portfolios. The last row (L-H) reports the spread return between low (undervalued) and high (overvalued) *VRES* firms. All various returns (with one exception) are

[14] Our results are similar when portfolios are formed annually at the end of June and held for the subsequent twelve months.

monotonically decreasing with higher levels of *VRES*. Over the sample period of July 1990 to June 2021, the L-H spread return amounts to a highly significantly and economically important 0.56% per month (i.e., 6.72% annually) after controlling for firm size and is in line with results documented in Nichols et al. (2017) and Walkshäusl (2021). The average value-weighted return is 0.30% per month and lower than the equal-weighted return of 0.83% which indicates that the return premium is more pronounced among smaller firms.

Table 3.4: Portfolio sorts based on VRES, June 1990 - June 2021.

Panel A reports average monthly returns in percent for quintile portfolios sorted on VRES. We report equal-weighted, value-weighted, and size-adjusted returns for portfolios formed at the end of each calendar quarter based on *VRES*. The last column (Low-High) reports the spread return between low and high *VRES* firms allocated to quintile portfolios. Newey and West (1987) corrected t-statistics are provided in parenthesis and */**/*** indicate significance at the 10%/5%/1% level. Panel B reports average firm characteristics for the portfolios as measured at the portfolio formation date (see the main text for details on variable definitions).

Panel A: Monthly return characteristics						
Return	(Low)	2	3	4	(High)	L-H
Equal-weighted	1.40^{***}	0.99^{***}	0.77^{***}	0.65^{**}	0.58^{**}	0.83^{***}
	(3.52)	(3.19)	(2.78)	(2.46)	(2.07)	(5.12)
Value-weighted	0.97^{***}	0.80^{***}	0.68^{***}	0.54^{**}	0.67^{***}	0.30^{**}
	(3.03)	(2.89)	(2.78)	(2.21)	(2.59)	(2.01)
Size-adj.	0.28^{***}	0.06	-0.09	-0.20^{***}	-0.28^{***}	0.56^{***}
	(3.58)	(1.62)	(-1.87)	(-4.00)	(-5.28)	(5.18)
Panel B: Average portfolio characteristics						
Variable	(Low)	2	3	4	(High)	
$VRES$	-7.72	-0.96	-0.26	0.13	0.58	
$SIZE$	633.02	1,537.05	2,373.93	2,998.17	3,076.93	
$BMRATIO$	1.21	1.10	0.92	0.76	0.64	
INV	0.13	0.11	0.11	0.12	0.16	
OP	0.72	0.70	0.71	0.76	0.84	
MOM	0.08	0.09	0.10	0.12	0.16	
REV	0.00	0.01	0.01	0.01	0.01	
ACC	-0.07	-0.05	-0.03	-0.04	-0.05	
NSI	0.11	0.07	0.06	0.07	0.09	
$\Delta XFIN$	0.09	0.03	0.02	0.02	0.05	
$FSCORE$	5.19	5.76	5.92	5.91	5.73	
$IVOL$	0.37	0.28	0.26	0.25	0.28	
$SKEW$	0.19	0.17	0.17	0.17	0.17	
MAX	0.07	0.05	0.05	0.05	0.05	

Fig. 3.1 illustrates the cumulative payoff of a $1.00 investment in the portfolio of low

(solid line) and high (dashed line) *VRES* stocks. The low *VRES* portfolio grows over time to \$9.86, whereas an investment in the overvalued, high *VRES* quintile portfolio only grows to \$3.85 for each dollar invested. In comparison, an investment in the global market portfolio comprising equities from developed countries results in a terminal value of \$6.03 until June 2021.

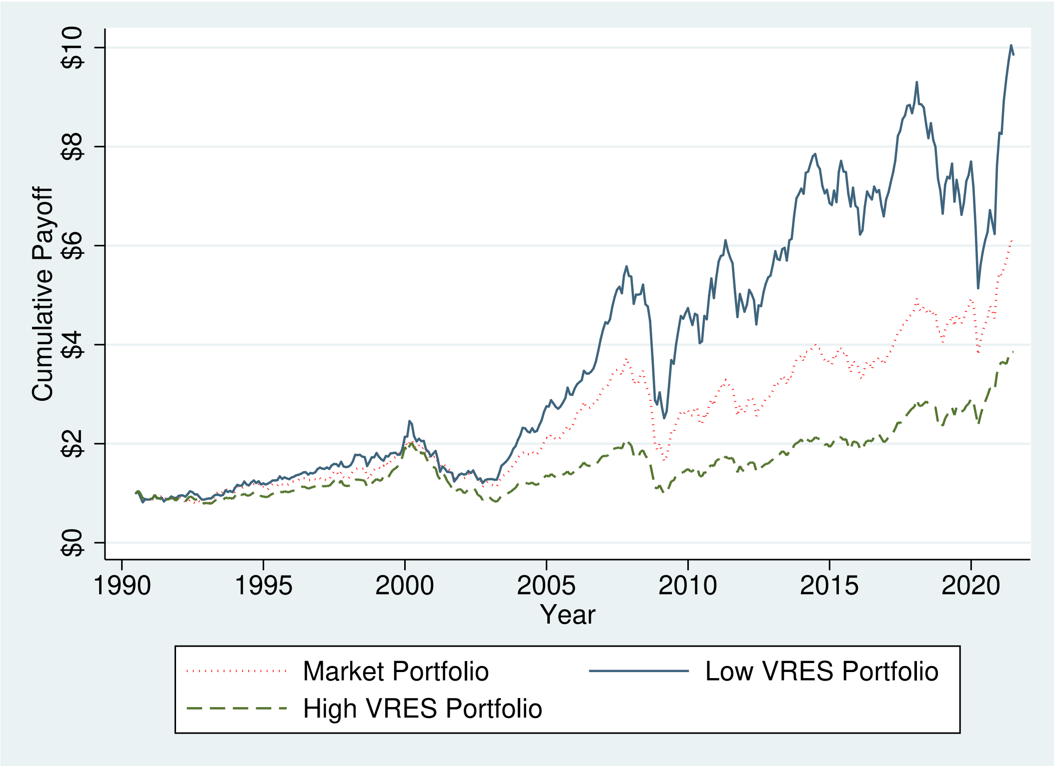

Fig. 3.1. Cumulative payoff of low (solid line) and high (dashed line) *VRES* stocks of a \$1.00 investment in the according quintile portfolio over the sample period. The dotted line shows an investment in the global market portfolio comprising all stocks from developed countries.

The profit of a portfolio strategy based on a firm-level proxy for mispricing should vary with the degree of market-wide mispricing. To be specific, a mispricing related investment strategy should generate higher returns when there is a higher degree of market-wide mispricing. To test this hypothesis, we measure market-wide mispricing as in Nichols et al. (2017) and use the adj. R^2 from our quarterly cross-sectional regressions in Eq. (3.3) net of its rolling 10-year average. For the initial years in our sample, the rolling average is calculated over the available observations.

Fig. 3.2 shows times of high (low) mispricing as below-average (above-average) adj. R^2 values. We detect an average monthly size-adj. L-H return of 0.67% (t-statistic: 4.47)

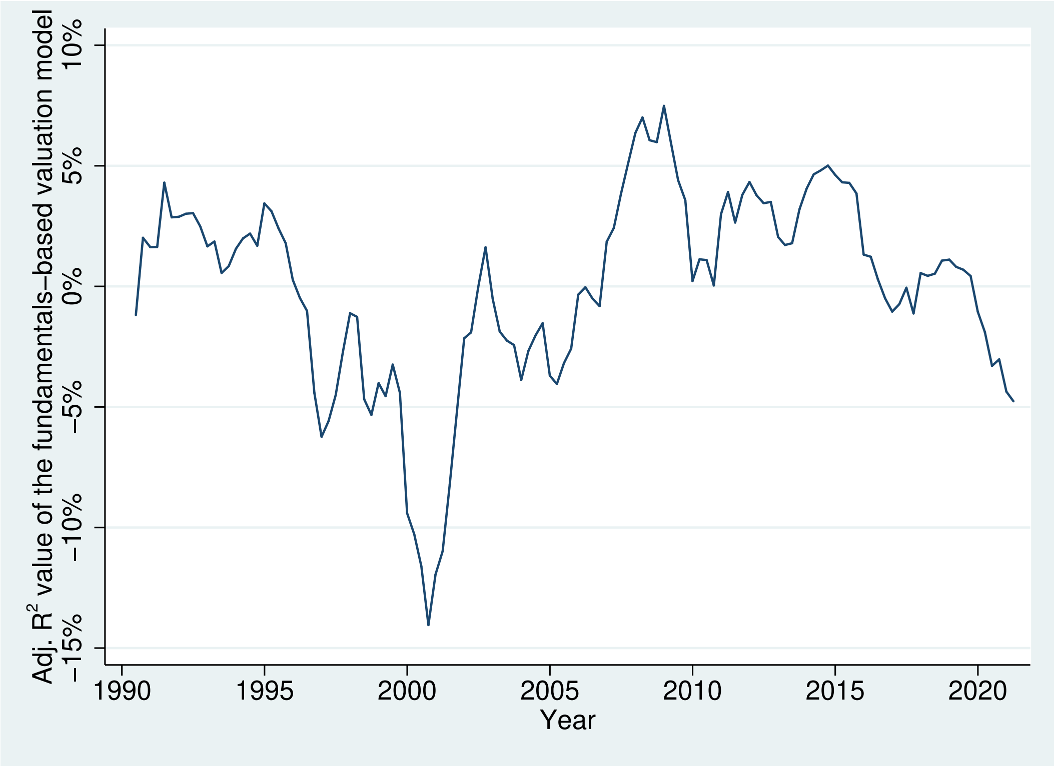

Fig. 3.2. This figure shows the adj. R^2 value of our extended fundamentals-based valuation model net of its rolling 10-year average. For early years, the rolling average is calculated over the available observations. Times of high (low) mispricing are indicated by below-average (above-average) values.

in times of high mispricing and 0.48% (t-statistic: 3.21) in times of low mispricing. September 2000 is identified as the month of highest mispricing in our sample which is only six months after the Nasdaq Composite peak on March 10, 2000, and coincides with the beginning of the subsequent dot-com stock market crash. As expected, the average size-adj. L-H return following the twelve months after September 2000 is a much higher than average 0.93% (t-statistic: 4.02). Taken together, the fundamentals valuation model not only provides evidence to capture firm-level mispricing, but also correctly identifies overall market-wide times of mispricing because of weaker mapping (i.e., lower values for the adj. R^2) of firm fundamentals into stock prices.

Our average portfolio characteristics shown in Panel B of Table 3.4 reveal that sorting stocks on *VRES* is related to some well-known determinants of future stock returns. Undervalued firms usually tend to be much smaller in terms of market capitalization ($633.02m) compared with overvalued firms ($3.08b) and their book-to-market ratio is almost twice as high (1.21, resp., 0.64). Similarly, their operating profitability (0.72) is relatively weaker and overvalued stocks generated a cumulative stock return of 16% within the twelve-month prior to our portfolio formation date, vastly outperforming undervalued stocks that only gained 8% on average. Besides a higher level of *IVOL* for undervalued firms (0.37, resp., 0.28 for high *VRES* firms), there are no economically notable differences in firm characteristics.

Our comprehensive sample of firms from global developed countries steps into the general debate whether portfolios should be built using global, regional, or local *VRES* breakpoints.[15] As a robustness test, we replicate our portfolio strategy and use both local, i.e., country specific sorts, and regional (Europe, Japan, and Asia-Pacific) sorts. Applying regional sorts, all results remain virtually unchanged. The size-adj. L-H return is a highly significant 0.60% per month and the according value-weighted return 0.26% per month. Sorting stocks based on country-specific breakpoints gives a highly significant size-adj. L-H return of 0.62% per month.

Finally, we also analyze portfolios that are constructed using only firms located in a single country as a robustness test. We are restricted in these country-specific L-H (quintile) portfolios by limited firm-observations to ensure well diversified portfolios. Japan, the United Kingdom, France, and Germany comprise 66% of aggregated market capitalization in our entire sample and each country has on average at least 214 firm observations each month. We observe highly significant size-adj. L-H returns of 0.48% (Japan), 0.47% (UK), 0.48% (France), and 0.64% (Germany). A further analysis of these local, value-weighted L-H returns reveals highly significant alphas according to the regional Fama/French five-factor model including an additional momentum factor of 0.40% (Japan), 0.51% (UK), and 0.80% (Germany), while the alpha of 0.27% in France is statistically insignificant.

In summary, our extensive portfolio analysis provides evidence that the value residual *VRES* is negatively related with stock returns in capital markets around the world.

3.4.2 Regression analysis

To examine the robustness of our portfolio analysis, we conduct cross-sectional regressions using the methodology in Fama and MacBeth (1973) in this section:

[15] See e.g., Griffin (2002), Chaieb et al. (2011), and Hollstein (2022) who prefer local, i.e., country level, variable breakpoints. Brooks and Del Negro (2005), Fama and French (2012), and Fama and French (2017) are, however, in favor of regional analysis.

$$
\begin{aligned}
r^e_{i,t+1} = {} & \gamma_{0,t} + \gamma_{1,t} VRES_{i,t} + \gamma_{2,t} \ln\left(SIZE_{i,t}\right) + \gamma_{3,t} BMRATIO_{i,t} + \gamma_{4,t} OP_{i,t} \\
& + \gamma_{5,t} INV_{i,t} + \gamma_{6,t} MOM_{i,t} + \sum_{j=7}^{J} \gamma_{j,t} CONTROL^j_{i,t} + \text{Country Dummies}_{i,t} \\
& + \epsilon_{i,t+1}.
\end{aligned} \tag{3.4}
$$

Formally, we regress firm-level one-month ahead stock returns minus the one-month Treasury bill rate, $r^e_{i,t+1}$, on *VRES*, common control variables, and unreported country dummies. OLS-regressions imply equal weights on all observations, thus emphasizing small and economically less important firms. To examine the economic strength of *VRES* across different levels of firms size, we mostly apply value-weighted regressions using the natural log of firm size as observation weights.[16] Coefficient estimates in cross-sectional regressions are, however, not easy to interpret. As an additional robustness test, we use an indicator variable *QVRES* whose regression estimates reflect a low-high portfolio strategy as analyzed in Section 3.4.1. To construct *QVRES*, we follow Bradshaw et al. (2006) and allocate all stocks in ascending order to quintile portfolios based on *VRES*. The indicator variable *QVRES* is a firms' numerical rank based on these quintile portfolio allocations, where the ranks are scaled to an interval between zero and one (formally, 1-[(5 - quintile rank)/4]). Thus, *QVRES* is zero (one) for firms in the bottom (top) *VRES* quintile portfolio which reflects the L-H portfolio strategy and results can be interpreted accordingly.

Table 3.5 supplements our portfolio analysis presented in Table 3.4 and shows that the value residual has a highly significantly explanatory power for future stock returns in all regression specifications. Controlling for Fama-French five-factor model related variables and momentum in Spec. (1), we observe an estimated coefficient for *VRES* of -0.05 with a robust t-statistic of -7.52. The negative sign is in line with our previous finding that low *VRES* firms offer a return premium for being undervalued with regards to accounting fundamentals. Otherwise, international stock returns are strongly related with book-to-market ratio, operating profitability, momentum, and investment, but not with firm size which is in line with recent studies on international stock returns (Fama and French (2017)).

[16] Using (equal-weighted) OLS-regressions or size-adjusted returns based on quintile size-portfolios as $r^e_{i,t+1}$ does not affect our main findings.

Table 3.5: Monthly cross-sectional return regressions (1990-2021) on VRES and common controls.

This table shows time-series averages for the coefficients from firm-level cross-sectional regressions of monthly stock excess returns on *VRES* and common control variables. *QVRES* scales the quintile rank of *VRES* portfolio allocations to range between zero and one. Formally, *QVRES* equals $1 - [(5 - \text{quintile rank})/4]$, and takes the value zero (one) for firms in the bottom (top) *VRES* quintile portfolio. The set of controls in all specifications includes the natural log of firm size (*SIZE*), book-to-market ratio (*BMRATIO*), operating profitability (*OP*), investment (*INV*), and momentum (*MOM*). Further controls are accruals (*ACC*) and net stock issues (*NSI*) as defined in Fama and French (2008b). Mispricing related controls include $\Delta XFIN$ (Bradshaw et al. (2006)) and *FSCORE* (Piotroski (2000)). To capture investor preferences for lottery-like stocks (Ang et al. (2006); Ang et al. (2009); Bali et al. (2011)), we control for the maximum (*MAX*), total skewness (*SKEW*), and idiosyncratic volatility relative to the Fama-Fremch five-factor model (*IVOL*), all of daily stock returns over the previous month. *REV* is the stock return over the previous month (Jegadeesh (1990)). Variables using accounting data as well as *SIZE* are updated at the end of June each year and *VRES* is updated each quarter. All other explanatory variables are updated each month. Country dummies are included in all specifications. The R^2 values are adjusted for degrees of freedom. *N* denotes the average number of firms and *T* the number of months. VW indicates that cross-sectional return observations are weighted by the natural log of firm size and EQ referrs to equal-weighted, standard OLS-regressions. *Period* indicates if coefficients are estimated using the full sample period (1990-2021) or sub-periods of low, resp. high, market-wide mispricing. Following Nichols et al. (2017), times of low (high) market-wide mispricing are identified by above-average (below-average) values of the adj. R^2 from the fundamentals-based valuation model net of its rolling 10-year average. Newey and West (1987) robust t-statistics are given in parenthesis and */**/*** indicate significance at the 10%/5%/1% level.

Specification	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VRES	−0.05***				−0.05***	−0.06***	−0.06***	−0.06***	−0.06***
	(−7.52)				(−7.71)	(−8.28)	(−8.33)	(−8.30)	(−8.95)
QVRES		−0.62***	−0.54***	−0.72***					
		(−9.33)	(−6.54)	(−6.71)					
ln(SIZE)	0.00	−0.01	0.01	−0.05	0.00	−0.01	−0.01	0.00	−0.01
	(0.04)	(−0.47)	(0.48)	(−1.09)	(0.00)	(−0.44)	(−0.38)	(0.07)	(−0.29)
BMRATIO	0.25***	0.21***	0.18***	0.26**	0.24***	0.24***	0.24***	0.25***	0.24***
	(4.05)	(3.53)	(3.06)	(2.25)	(3.98)	(3.93)	(3.93)	(4.17)	(4.04)
OP	0.11***	0.12***	0.11***	0.12**	0.11***	0.09**	0.10***	0.10***	0.09***
	(4.06)	(4.04)	(3.15)	(2.51)	(4.17)	(3.56)	(3.90)	(3.78)	(3.37)
INV	−0.55***	−0.56***	−0.72***	−0.34**	−0.50***	−0.24***	−0.54***	−0.56***	−0.22***
	(−5.42)	(−5.38)	(−5.50)	(−2.39)	(−5.36)	(−2.81)	(−5.51)	(−5.59)	(−2.65)
MOM	0.78***	0.77***	0.57	1.03***	0.77***	0.75***	0.81***	0.79***	0.76***
	(3.04)	(3.04)	(1.46)	(3.57)	(3.02)	(2.98)	(3.36)	(3.23)	(3.14)
ACC					−0.08*				−0.09**
					(−1.83)				(−2.07)
NSI					−0.21**				−0.16**
					(−2.37)				(−2.11)
$\Delta XFIN$						−0.70***			−0.68***
						(−3.82)			(−4.01)
FSCORE						0.05***			0.05***
						(3.60)			(4.12)
IVOL							−0.28	0.99***	1.10***
							(−1.04)	(2.98)	(3.42)
SKEW							−0.03	0.18***	0.18***
							(−1.03)	(5.75)	(5.79)
MAX								−3.63***	−3.10***
								(−2.83)	(−2.52)
REV								−2.10***	−2.22***
								(−4.29)	(−4.57)
Adj. R^2	0.10	0.10	0.09	0.10	0.10	0.10	0.10	0.11	0.11
N	3,797	3,797	4,008	3,524	3,797	3,797	3,797	3,797	3,797
Weighting	VW	EQ	EQ	EQ	VW	VW	VW	VW	VW
T	372	372	210	162	372	372	372	372	372
Period	Full	Full	Low Mis.	High Mis.	Full	Full	Full	Full	Full

Results using the indicator variable *QVRES* in Spec. (2) shows that the abnormal return difference between the lowest and highest quintile portfolio is a highly significant -0.62% per month after controlling for size, book-to-market ratio, operating profitability, investment, and momentum.[17] After having analyzed the full sample period, the key question is does the according return premium earned by a *VRES* related investment strategy vary over time? In addition to our findings for portfolio strategies in Section 3.4.1, Spec. (3) and (4) in Table 3.5 include only months that are identified as times of low, resp. high, market-wide mispricing based on the net of rolling 10-year average adj. R^2 values from the fundamentals-based valuation model as stated in Eq. (3.3). We observe an abnormal return difference of -0.54% in times of low mispricing and the spread return is larger in magnitude in times of high mispricing (-0.72%). Our coefficient estimates for *QVRES* can be interpreted as portfolio return differences and are generally in the same order of magnitude as the results presented in Section 3.4.1. In consequence, we document a robust and economically important cross-sectional relation between the value residual *VRES* and one-month ahead stock returns in our global analysis.

Controlling for other well-known cross-sectional predictors of stocks returns in Spec. (5) to (9) does not affect the high explanatory power of the value residual. In result, *VRES* captures firm-level aspects of mispricing not already identified by fundamental strength as proxied by *FSCORE* (Piotroski (2000)) or external financing activities as measured by $\Delta XFIN$ (Bradshaw et al. (2006)). Controlling for lottery-like stock preferences in Spec. (7) and (8) also does not affect the estimated coefficient for *VRES*.[18] Including the broad set of all control variables provides evidence that *VRES* captures information about future stock returns left unexplained by these predictors. Our regression results lead us to the conclusion that the economic effect of the value residual is at least a significant -0.50% per month (formally $-0.06 \times (0.58 - (-7.72))$, resp. 6% per year, between stocks in the lowest *VRES* quintile and stocks in the highest *VRES* quintile portfolio. This return difference, however, is not explained away in the presence of other return determinants.

[17] As a robustness test and to mitigate the influence of small stocks in cross-sectional regressions, we estimate Spec. (2) using value-weighted regression using the natural log of firm size as observation weights. The according abnormal return difference is -0.60% per month (t-statistic: -9.18).

[18] The fact that idiosyncratic volatility has a significantly positive relation with future stock returns in Spec. (8) is explained by contemporaneously controlling for *MAX* as documented in Bali et al. (2011) for the U.S. and in Cheon and Lee (2017) for international markets.

3.5 Robustness tests

Do investors who buy undervalued stocks as indicated by low values of *VRES* and sell high *VRES* stocks gain abnormal returns by exploiting mispricing opportunities or by facing exposure to risk? In this section, we test risk- and mispricing-based explanations and critically discuss the applied methodology as well as our main results. The first subsection focuses on the time-series analysis of portfolios sorted on *VRES*. The next subsection analyzes the dynamics of migrations among these portfolios, followed by robustness tests for empirical implications related with investor sentiment and arbitrage capital.

3.5.1 Time-series analysis, the post-publication effect, and asset pricing implications

We start our robustness tests at the portfolio level using *VRES* factor mimicking portfolios as proposed in Fama and French (1992) and Fama and French (1993). The construction of these long-short zero-cost portfolios in this section follows common principles for international markets as shown in Fama and French (2017). At the end of each quarter, we sort stocks within a region on market capitalization and denote those in the top 90% of cumulative market capitalization for that specific region as big stocks and those in the bottom 10% as small stocks. The *VRES* breakpoints in the 2×3 sorts for a region are the 30th and 70th percentiles of *VRES* for big stocks of the region. The intersection of the independent sorts on *SIZE* and *VRES* produces six portfolios and we compute monthly value-weighted returns for each over the next three months. The *VRES* factor mimicking portfolio for a region is the equal-weighted average of the returns for the two high *VRES* portfolios for a region minus the average of the returns for the two low *VRES* portfolios. We use these portfolio returns in time-series regressions on common risk factors:

$$\begin{aligned} r_t^P = \alpha &+ \gamma_1 PS_t + \gamma_2 PP_t + \gamma_3 MKTRF_t + \gamma_4 SMB_t \\ &+ \gamma_5 HML_t + \gamma_6 CMA_t + \gamma_7 RMW_t + \gamma_8 MOM_t + \epsilon_t. \end{aligned} \tag{3.5}$$

r_t^P denotes the monthly value-weighted return of either the long leg, short leg, or their according return difference, of our *VRES* factor mimicking portfolios. *PS* (post sample) is a dummy variable equal to one if month t is after the end of the original sample period in Nichols et al. (2017) but still pre-publication, and zero otherwise. Similarly, *PP* (post publication) is a dummy variable equal to one if month t is after publication of the original study and zero otherwise. *MKTRF* is the market excess return, *SMB*, *HML*, *CMA*, *RMW*, and *MOM*, are, respectively, factor returns for the developed ex-U.S. region provided by Kenneth French. Because of limited data availability for the momentum factor, the sample period in this section starts in Nov. 1990.

Eq. (3.5) disentangles whether the observed return difference among varying levels of *VRES* is more consistent with a risk-based or mispricing-based interpretation in two ways. First, our concerns that low *VRES* firms only earn higher returns because they are on average small, value firms, are accounted for by controlling for SMB and HML. Second, a mispricing-based explanation is supported if academic research draws trading attention to exploitable arbitrage opportunities which should be reflected in a decline of abnormal returns after publication (see e.g., Hanson and Sunderam (2014); McLean and Pontiff (2016)).[19] The original study by Nichols et al. (2017) was published on May 26, 2017, and to the best of our knowledge, no preprint circulated among academics or practitioners.[20] For that reason, we set *PS* equal to the value of one for all months from January 2013 to May 2015 and *PP* equals one for all subsequent months.

Panel A of Table 3.6 shows that the average value-weighted return for the global *VRES* portfolio strategy is a highly significant 0.29% per month. On the regional level, we observe slightly higher average returns of 0.40% in Japan and smaller returns of 0.26% in Europe, whereas the average return in Asia-Pacific is an insignificant 0.20%. Interestingly, the short portfolios earn positive returns in all regions which is surprising, because it economically implies that a potential overvaluation is not corrected in form of subsequent negative, abnormal returns.[21] At least for the period between 2018 and 2020 not yet analyzed in previous studies, this finding is in line with Blitz (2021) who shows that only the largest and most expensive growth stocks outperformed the market return in developed countries (incl. the U.S.). An in depth analysis of the underlying 2×3 subportfolios reveals that it is indeed the high return of 0.65% for the high *VRES* portfolio among big firms that significantly lowers our overall L-H spread returns compared with Nichols et al. (2017) and Walkshäusl (2021).

[19] We further address limitations to arbitrage later in Section 3.5.5.

[20] The commonly used preprint archive SSRN does not offer a working paper version of their study and a similar search on Google Scholar only refers to the published version in *Contemporary Accounting Research*.

[21] At least within the next three months after which our portfolios are rebalanced.

Table 3.6: *VRES* factor mimicking portfolio return analysis.

Panel A reports average monthly returns in percent for *VRES* factor mimicking portfolios based on the intersection of independent bivariate sorts on firm size and *VRES*. At the end of each quarter, we sort stocks in a region (Europe, Japan, Asia-Pacific, or global) in two size groups and three *VRES* groups. Big stocks are separated from small stocks by the level of top 90% cumulative market capitalization for a region. The *VRES* breakpoints are the 30th and 70th percentiles of *VRES* for big stocks in a region. We compute monthly value-weighted returns for each of the 2×3 portfolios over the next three months. The *VRES* factor mimicking portfolio for a region is the equal-weighted average of the returns for the two high *VRES* portfolios for a region minus the average of the returns for the two low *VRES* portfolios. Panel B shows time-series regression estimates for our *VRES* factor mimicking portfolio returns on regional risk factors provided by Kenneth French and on two dummy variables. *Post Sample* is a dummy variable equal to one if month t is after the end of the original sample period in Nichols et al. (2017) but still pre-publication, and zero otherwise. *Post Publication* is a dummy variable equal to one if month t is after publication of the original study and zero otherwise. $MKTRF$ is the value-weighted market return minus the U.S. one-month Treasury bill rate. *SMB*, *HML*, *RMW*, *CMA*, and *MOM* are as defined in Fama and French (2018). Newey and West (1987) corrected t-statistics are given in parenthesis and */**/*** indicate significance at the 10%/5%/1% level. For the Fama-French five factor model incl. momentum (base model) and an augmented model (VRES) including the *VRES* factor return, Panel C shows the actual, ex-post maximum squared Sharpe ratio Sh^2 and out-of-sample mean and median estimates $\hat{Sh}^2$ using the bootstrap approach described in Fama and French (2018) with 100,000 runs. The last three columns of Panel C report the differences of out-of-sample estimated mean and median values $\hat{Sh}^2_{VRES} - \hat{Sh}^2_{base}$ and the percentage of simulation runs in which this difference is negative.

Panel A: Average monthly value-weighted returns (July 1990 - June 2021)

	Europe			Japan			Asia-Pacific			Global		
	Short	Long	L-S	Short	Long	L-S	Short	Long	L-S	Short	Long	L-S
	0.72**	0.98***	0.26***	0.21	0.60**	0.40***	0.84**	1.04***	0.20	0.58**	0.87***	0.29***
	(2.54)	(3.17)	(2.93)	(0.67)	(2.07)	(3.19)	(2.27)	(2.76)	(1.34)	(2.26)	(3.20)	(3.46)

Panel B: Time-series regressions on common risk factors (Nov. 1990 - June 2021)

	Europe			Japan			Asia-Pacific			Global		
	Short	Long	L-S	Short	Long	L-S	Short	Long	L-S	Short	Long	L-S
Alpha	0.19**	0.49***	0.30**	0.11	0.41***	0.31***	0.30**	0.46***	0.16	0.17	0.43***	0.26***
	(2.10)	(6.68)	(2.54)	(1.03)	(4.90)	(2.90)	(1.97)	(3.72)	(1.02)	(1.62)	(4.77)	(2.73)
Post Sample	−0.30**	−0.49***	−0.19	−0.58***	−0.16	0.42**	−0.11	−0.45**	−0.34	−0.33***	−0.32***	0.01
	(−2.13)	(−4.51)	(−0.88)	(−2.90)	(−1.33)	(2.07)	(−0.42)	(−1.97)	(−1.00)	(−2.73)	(−3.12)	(0.04)
Post Publication	0.13	−0.41***	−0.53***	−0.16	−0.21	−0.04	0.44	−0.21	−0.65	0.04	−0.39***	−0.43**
	(1.09)	(−3.32)	(−2.88)	(−0.80)	(−1.58)	(−0.17)	(0.97)	(−0.99)	(−1.15)	(0.20)	(−3.79)	(−2.19)
MKTRF	1.02***	1.04***	0.02	1.00***	0.95***	−0.05*	0.96***	0.98***	0.02	0.92***	0.97***	0.05***
	(51.28)	(57.26)	(0.61)	(55.19)	(58.96)	(−1.91)	(32.87)	(37.81)	(0.55)	(54.15)	(62.46)	(2.63)
SMB	0.34***	0.42***	0.08	0.30***	0.29***	−0.02	0.37***	0.50***	0.13**	0.35***	0.35***	0.00
	(8.85)	(17.21)	(1.58)	(7.25)	(10.03)	(−0.40)	(8.23)	(14.15)	(2.43)	(7.75)	(9.77)	(0.01)
HML	−0.19***	0.06	0.25***	−0.11*	0.21***	0.32***	−0.20***	0.07	0.27***	−0.02	0.23***	0.25***
	(−4.85)	(1.34)	(4.73)	(−2.03)	(5.70)	(4.56)	(−2.89)	(1.29)	(3.07)	(−0.31)	(4.73)	(4.14)
RMW	0.10*	0.19***	0.09	0.27***	0.27***	0.00	−0.13*	0.05	0.18*	0.22**	0.31***	0.08
	(1.86)	(3.95)	(1.51)	(2.68)	(4.58)	(0.01)	(−1.65)	(0.77)	(1.85)	(2.54)	(3.77)	(1.16)
CMA	0.13***	0.05	−0.08	0.19**	0.13***	−0.06	0.05	0.06	0.02	−0.11	−0.13**	−0.02
	(2.71)	(0.76)	(−0.84)	(2.29)	(2.62)	(−0.66)	(0.58)	(0.91)	(0.21)	(−1.49)	(−2.00)	(−0.26)
MOM	−0.04	−0.08***	−0.04	−0.04	−0.12***	−0.08**	−0.06	−0.02	0.04	−0.01	−0.05**	−0.04
	(−1.42)	(−3.93)	(−1.10)	(−1.14)	(−5.43)	(−1.98)	(−1.63)	(−0.55)	(1.07)	(−0.23)	(−2.49)	(−0.90)
Adj.R^2	0.95	0.97	0.17	0.91	0.94	0.22	0.88	0.92	0.07	0.93	0.95	0.19

Panel C: Comparison of asset pricing models via bootstrapping approach in Fama and French (2018)

	FF5+MOM			FF5+MOM+VRES			Out-of-sample difference		
Region	Actual	Mean	Median	Actual	Mean	Median	Mean	Median	% < 0
Europe	0.2161	0.1681	0.1556	0.2335	0.1750	0.1621	0.0069	0.0055	37.11
Japan	0.0319	0.0169	0.0090	0.0673	0.0372	0.0277	0.0203	0.0143	15.82
Asia-Pacific	0.2465	0.1873	0.1768	0.2466	0.1785	0.1680	−0.0088	−0.0027	66.52
Global	0.2304	0.1863	0.1746	0.2560	0.2002	0.1883	0.0138	0.0123	29.58

Coefficient estimates reported in Panel B of Table 3.6 confirm that returns of our *VRES* related portfolio strategy strongly covary with the value factor. We find a highly significant intercept (alpha) of 0.29% for our global portfolio strategy return left unexplained by size, value, profitability, investment, and momentum factor returns. Similarly, we find highly significant intercepts in Europe (0.26%) and Japan (0.40%), with the exception of Asia-Pacific (insignificant 0.20%). The most interesting aspect in Panel B is the significant negative coefficient estimate for our post publication dummy *PP* in Europe and the global sample. These results are in favor of a mispricing explanation where academic research draws attention to potential arbitrage opportunities. This argument is further strengthened by the fact that the post publication decline is only prevalent among the long portfolio leg which comprises undervalued stocks. In a market with mispriced stocks, investors can very easily exploit undervaluation by taking a long position while they may be unwilling to short sell overvalued stocks because of institutional constraints, short sale impediments, or an inherent risk of arbitrage as discussed in Miller (1977) or Shleifer and Vishny (1997).

Interestingly, Jacobs and Müller (2020) finds that the U.S. is the only market with a reliable post publication decline among 39 analyzed stock markets. Based on their taxonomy, *VRES* is a valuation predictor, and they report a significant decline in value-weighted returns of 0.45% for similar anomalies in the U.S. but an insignificant decline in international markets. In perspective, we observe a decline of 0.43% for the global *VRES* hedge portfolio which stems nearly entirely from a return decline of 0.39% within its long leg. The p-value of a Wald test with the null hypothesis that the sum of α, PS, and PP is jointly zero is 0.46, thus not rejecting the null. In other words, the entire return left unexplained by common risk factors disappears after publication of the original study in May 2017, indicating that an according portfolio strategy actually exploits mispricing and according arbitrage opportunities.

In Panel C of Table 3.6, we further evaluate the degree of *VRES* to explain average stock returns by a right-hand-side (RHS) approach proposed in Fama and French (2018). We compare the global, developed ex-U.S. Fama-French five-factor model incl. momentum with a nested model including the global *VRES* factor returns using the actual maximum squared Sharpe ratio Sh^2 in addition to mean and median estimates $\hat{Sh}^2$ from out-of-sample bootstrap simulations. According to Fama and French (2018), we split our sample of $T = 368$ month (Nov. 1990 - June 2021) into $T/2$ adjacent pairs $(1, 2)$,$(3, 4)$,...$(T-1, T)$. In each of our 100,000 simulation runs, we draw $T/2$ pairs with replacement and randomly assign a month from each pair to the in-sample subset, and the other month from each pair to the out-of-sample subset. The assignments are taken from the first occurrence if a pair is drawn more than once. Next, for both asset pricing models and in each simulation run we identify weights of the included factors resulting in maximizing Sh^2 for all in-sample month. To finally compute out-of-sample $\hat{Sh}^2$ estimates, we combine the in-sample weights with out-of-sample means and covariances of the respective factors. We report mean and median for each of the 100,000 simulation runs, as well as the according differences between our two considered models. The last column of Panel C shows the percentage rate of $\hat{Sh}^2$ estimates for the global, developed ex-U.S. Fama-French five-factor model incl. momentum having values less than $\hat{Sh}^2$ estimates with an augmented *VRES* factor.

The actual squared Sharpe ratio Sh^2 of the developed ex-U.S. Fama-French five-factor model incl. momentum is 0.2304 with an out-of-sample mean of 0.1863. Most of the weights of the Sharpe ratio maximizing portfolio is on the profitability factor (46%) and the value factor (17%).[22] We observe a squared Sharpe ratio of 0.2560 for the *VRES* augmented model, thus an increase of 0.0256. To put this in perspective for ease of interpretation, additionally including *VRES* to the base model contributes to its explanatory power approximately in the same magnitude as the momentum factor adds to the U.S. Fama-French five-factor model within the period July 1963 to June 2016 (see Table 5 in Fama and French (2018)).[23] The Sh^2 of the augmented model exceeds Sh^2 of the developed ex-U.S. Fama-French five-factor model incl. momentum in more than 70% of our simulation runs. Overall, *VRES* expands the investment opportunity set already trading the market portfolio and well established risk factors in all regions except Asia-Pacific.

Risk-based asset pricing models compete with mispricing related models that also use a parsimonious set of factors to explain differences in expected returns. Stambaugh and Yuan (2017) propose a model based on two factors, *PERF* and *MGMT*, that can loosely be interpreted as short-term and long-term mispricing factors. Together with a market and size factor, the model aims to caputure time-variation in common mispricing. We thankfully use data for this mispricing model from Hanauer (2020) who provides global ex-U.S. factor returns beginning in July 1990. Over our entire sample period, the average return of *PERF* is 0.44% and the return of *MGMT* is 0.22%, both highly significantly different from zero.

[22]The low weight of only 11% for the market portfolio, surprising at first, is explained in Lochstoer and Tetlock (2020).

[23]More formally, the contribution of momentum is 0.023/0.190.

Table 3.7: Monthly time-series regressions, July 1990 - June 2021.

This table shows time-series regression estimates for our *VRES* factor mimicking portfolio returns on international, non-U.S. factors for the mispricing model proposed in Stambaugh and Yuan (2017). Factor returns are provided by Hanauer (2020). $MKTRF$ is the value-weighted market return minus the U.S. one-month Treasury bill rate. *SMB*, *PERF*, and *MGMT* are, respectively, factor returns for size, short-term mispricing, and long-term mispricing, as defined in Stambaugh and Yuan (2017). *Post Sample* is a dummy variable equal to one if month t is after the end of the original sample period in Nichols et al. (2017) but still pre-publication, and zero otherwise. *Post Publication* is a dummy variable equal to one if month t is after publication of the original study and zero otherwise. Newey and West (1987) corrected t-statistics are given in parenthesis and */**/*** indicate significance at the 10%/5%/1% level.

	Europe	Japan	Asia-Pacific	Global
Alpha	0.41***	0.42***	0.29*	0.37***
	(4.07)	(3.12)	(1.90)	(4.47)
Post Sample	−0.29	0.22	−0.39	−0.11
	(−1.15)	(0.73)	(−1.10)	(−0.70)
Post Publication	−0.69***	−0.15	−0.69	−0.57**
	(−2.91)	(−0.42)	(−1.20)	(−2.22)
MKTRF	0.03	−0.11**	0.01	0.04**
	(1.07)	(−2.51)	(0.30)	(2.08)
SMB	0.04	−0.02	0.17**	0.02
	(0.79)	(−0.36)	(2.10)	(0.50)
PERF	−0.06	−0.10	−0.00	−0.06
	(−1.00)	(−1.02)	(−0.04)	(−0.99)
MGMT	−0.02	0.24	0.19	0.09
	(−0.31)	(1.65)	(1.41)	(0.94)
$Adj.R^2$	0.03	0.08	0.02	0.06

Table 3.7 shows coefficient estimates for time-series regressions of our regional and global *VRES* factor mimicking portfolio returns r_t^{VRES} on the Stambaugh and Yuan (2017) model factor returns including post sample and post publication dummies:

$$r_t^{VRES} = \alpha + \gamma_1 PS_t + \gamma_2 PP_t + \gamma_3 MKTRF_t + \gamma_4 SMB_t + \gamma_5 PERF_t + \gamma_6 MGMT_t + \epsilon_t. \quad (3.6)$$

Estimated slopes on dummy variables, *MKTRF*, and *SMB* are nearly identical to our former presented results. The same holds for our highly significant alphas of 0.37% for the global *VRES* related strategy, resp. 0.41% in Europe and 0.42% in Japan. Again, the alpha in Asia-Pacific is a positive 0.29% but statistically only weakly significant. However, all intercepts are larger in magnitude compared with previous results from Table 3.6 which is very surprising because the factor model of Stambaugh and Yuan

(2017) is deliberately constructed to capture common effects of mispricing. However, in all regions and globally, estimated slope coefficients on both mispricing factors are statistically insignificant and their magnitude is too small to be of economic importance. In short, there is no evidence that returns of our *VRES* portfolio strategy are driven by exposures to clusters of the 11 anomalies in Stambaugh and Yuan (2017) underlying the construction of the mispricing factors *PERF* and *MGMT*.

For further robustness of our results, we also apply the behavioral factor model developed by Daniel et al. (2020a), although according data is only available for the U.S. and only until December 2018.[24] In line with former results, we observe a highly significant alpha of 0.39% (t-statistic: 4.27). Both proposed factors, *PEAD* (short-horizon mispricing) and *FIN* (long-horizon mispricing), are, however, statistically only weakly related with our *VRES* portfolio returns.

Overall, low *VRES* (i.e., undervalued) firms earn higher subsequent stocks returns than high *VRES* (i.e., overvalued) firms around the world. The according return premium is not subsumed by established cross-sectional return determinants and is not captured by already known forms of mispricing commonality.

3.5.2 Portfolio migration

Another way to evaluate if *VRES* related returns are associated with either risk or mispricing is to look at portfolio transitions as in Fama and French (2007). If *VRES* sorts firms primarily on risk characteristics, returns of extreme quintile *VRES* portfolios should be driven by stocks remaining in these portfolio. Otherwise, if mostly driven by leaving stocks, a risk based explanation is less likely because this means that the risk implied either reverses, resp. decays, or changes very quickly over the period of only three months.[25]

The left Panel of Fig. 3.3 plots the average cumulative size-adj. payoff of $1.00 invested in the lowest *VRES* quintile portfolio twelve months prior to the portfolio formation month. As expected, stocks that are labeled as undervalued and enter the low *VRES*

[24]The use of U.S. factor returns to explain the returns of our global non-U.S. *VRES* portfolio strategy is motivated by empirical evidence that developed equity markets have become integrated over time (Eun and Lee (2010)) and that the U.S. market has a leading role for international equity markets (Rapach et al. (2013)).

[25]We present detailed results for portfolio migrations for the lowest *VRES* quintile portfolio for brevity. However, results for the highest quintile portfolio also support our main findings in this section which are in favor of a mispricing-related explanation: Stocks that remain in (leave) the highest *VRES* quintile experience a cumulative size-adj. return of 3.90% (-5.97%) over the year following the portfolio formation month.

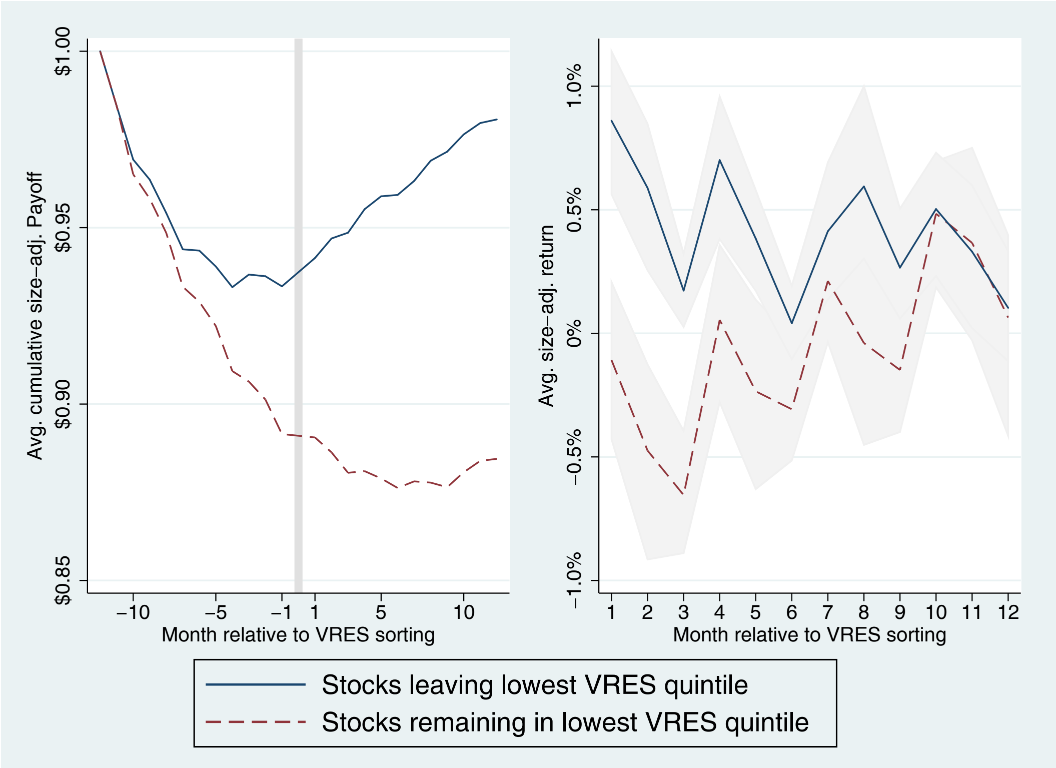

Fig. 3.3. This figure shows the average cumulative size-adjusted payoff of $1.00 invested in the lowest *VRES* quintile portfolio twelve months prior to the formation date up to twelve months after the formation date (left panel). The solid line referrs to the payoff of stocks that enter the lowest *VRES* quintile portfolio in t_0 and migrate out of that portfolio at the subsequent portfolio formation date. Similarily, the dashed line referres to the payoff of stocks that enter the lowest *VRES* quintile portfolio and remain for at least two formation dates (including t_0). The right panel shows average size-adjusted returns for both leaving and remaining firms for twelve months after the portfolio formation date. The shaded area covers the 95% confidence interval based on bootstrapped standard errors using the method of Politis and Romano (1994) with 100 simulation runs for each month and a block length of twelve months.

quintile portfolio underperformed firms with similar size by -8.76% over the previous year. However, differences arise between stocks that remain or leave the lowest quintile portfolio in subsequent periods. In line with a mispricing based explanation, stocks that transition out of the low *VRES* portfolio earn a cumulative size-adj. return of 5.07% over the year following the portfolio formation month, whereas remaining stocks perform poorly (-0.79%). In more detail, the right Panel of Fig. 3.3 shows the average size-adj. returns in each of the twelve months following the portfolio formation for the low *VRES* portfolio. We observe significant return differences between remaining and leaving stocks up to five months.[26] In the first month after portfolios are formed, leaving stocks earn

[26]The shaded area covers the 95% confidence interval based on bootstrapped standard errors using the method of Politis and Romano (1994) with 100 simulation runs for each month and a block length of twelve months.

size-adj. returns of 0.86% while remaining ones lose 0.11%. Within the three months until portfolios are updated, leaving (remaining) stocks generate a cumulative size-adj. return of 1.63% (-1.23%) and according differences are highly significant. Overall, our results are similar to Nichols et al. (2017) and more in favor of a mispricing-related explanation for the *VRES* investment strategy rather than a risk based explanation.

3.5.3 Investor sentiment

Investment strategies that are, at least partially, associated with the exploitation of mispricing are linked with market wide investor sentiment as shown in Stambaugh et al. (2012). Mispricing should be unlikely in form of underpricing in a market with some well-informed investors, so the primarily form is overpricing which is harder to correct if those investors are unwilling to sell short. In consequence, long-short portfolio returns should be higher following periods of high investor sentiment when many stocks are valued overly optimistic. To be more specific, the short leg of our portfolio strategy (consisting of overpriced stocks as indicated by high values of *VRES*) is expected to generate a greater profit following high sentiment while the long leg return should be less sensitive to investor sentiment.

To proxy for investor sentiment, we use the monthly (raw) index constructed by Baker and Wurgler (2006). Although the index is based on U.S. data, international investor sentiment is primarily driven by U.S. sentiment as shown in Baker et al. (2012), international market returns are closely linked to the lagged U.S. market return (Rapach et al. (2013)), and our sample of developed countries is a plausible set of integrated countries without market segmentations in the manner of Bekaert et al. (2007).

We determine the relation between our monthly global portfolio excess returns and sentiment effects with time-series regressions on one-month lagged levels of the sentiment index S_t:

$$
\begin{aligned}
r_t^{long} - rf_t &= \underset{(2.62)}{0.73} \underset{(-\,2.29)}{-\,0.83}\, S_{t-1} + \epsilon_t \\
r_t^{short} - rf_t &= \underset{(1.82)}{0.46} \underset{(-\,3.29)}{-\,1.09}\, S_{t-1} + \epsilon_t \\
r_t^{long-short} &= \underset{(3.21)}{0.27} + \underset{(2.00)}{0.25}\, S_{t-1} + \epsilon_t.
\end{aligned}
\tag{3.7}
$$

The estimated slope coefficients on both portfolio legs are negative, consistent with overall sentiment effects. The coefficient for the short leg is higher in magnitude, as expected, because this leg comprises overvalued stocks and they should be even more overpriced in times of high sentiment. A one-standard deviation increase in sentiment is associated with long-short returns increasing by 15 bps. We observe similar results at the regional level, with slopes for the short leg being larger in magnitude by 9% in Europe, 39% in Japan, and 73% in Asia-Pacific. The data hence support our mispricing related explanation for the *VRES* strategy in which market-wide sentiment creates overpricing due to short-sale limitations in the manner of Stambaugh et al. (2012).

Further, we test the hypothesis that anomalies, to the extent they reflect mispricing, should be stronger following high sentiment. A high-sentiment month is one in which the value of the sentiment index in the previous month is above the median value for our sample period, and the low-sentiment months are those with below-median values. In line with a mispricing explanation, our global *VRES* investment strategy earns on average a higher return of 0.41% (t-statistic: 4.34) in times of high sentiment compared with the average return of only 0.17% (t-statistic: 1.66) in times of low sentiment.[27]

3.5.4 Market-wide mispricing and anomaly return premiums

Time variation in expected risk premiums should be related to macroeconomic or market conditions according to risk-based explanations (see Chen et al. (1986) and Fama and French (1989) among others). Many attempts to link the return premiums for size, value, or momentum with sources of macroeconomic risks, however, have been quite unsuccessful (see e.g., Lakonishok et al. (1994), Asness et al. (2013), Bergbrant and Kelly (2016), or Koijen et al. (2018)). Others suggest a behavioral explanation, i.e., high returns are a manifestation of the correction of mispricing (see e.g., Rhodes–Kropf et al. (2005) or Jaffe et al. (2020)).

The analysis of the results shown in Table 3.5, specifications (3) and (4), suggests that risk controls vary with the level of market-wide mispricing. For example, momentum appears strong and significant during periods of high market mispricing but weakens when mispricing is low. This pattern holds for other variables like book-to-market, operating profitability, and investment, albeit with somewhat smaller effects across the two mispricing regimes. These findings align with the work of Walkshäusl (2016) in the U.S. equity market, where factors from the Fama-French five-factor model, especially HML (book-to-market), RMW (operating profitability), and CMA (investment), are linked to mispricing. To bolster our study's mispricing conclusions, we investigate whether

[27]These t-statistics are based on heteroskedasticity consistent standard errors of White (1980).

these findings extend to our global sample. To be specific, if the value residual is able to capture mispricing at the firm-level, an aggregated measure is expected to separate market-wide mispricing regimes. Based on the findings of Walkshäusl (2016), average factor premiums covary with the degree of market-wide mispricing and we should detect according return differences.

Table 3.8: Developed ex-U.S. factors, November 1990 - June 2021.

This table reports average monthly returns for developed ex-U.S. factors. *MKTRF* denotes the value-weighted market return minus the one-month Treasury bill rate. *SMB* (size), *HML* (value), RMW (operating profitability), *CMA* (investment), and *MOM* (momentum) are as defined in Fama and French (2018). Times of low (high) mispricing are identified by positive (negative) net of 10-year rolling average $adj.R^2$ values from cross-sectional regressions in our extented fundamentals-based valuation model. Newey and West (1987) corrected t-statistics are given in parenthesis and */**/*** indicate significance at the 10%/5%/1% level.

	MKTRF	SMB	HML	RMW	CMA	MOM
Full sample	0.39	0.10	0.28*	0.36***	0.11	0.65***
	(1.43)	(0.99)	(1.64)	(4.77)	(0.84)	(2.96)
Low Mispricing	0.21	0.01	0.03	0.34***	0.16	0.46*
	(0.55)	(0.07)	(0.25)	(3.88)	(1.21)	(1.77)
High Mispricing	0.64*	0.22	0.62*	0.37***	0.05	0.91**
	(1.67)	(1.17)	(1.77)	(2.81)	(0.22)	(2.39)

To separate times of high and low market-wide mispricing, we apply the net of 10-year rolling average adj. R^2 values of the fundamentals-based valuation model in Eq.(3.3) shown in Fig. 3.2. As summarized in Table 3.8, the developed ex-U.S. momentum factor on average earns a highly significant 0.65% per month. However, momentum returns are on average a weakly significant 0.46% during the 210 months of low market-wide mispricing but nearly twice as high in the 162 months of high mispricing. Even more pronounced, return premiums for the size and value anomaly are entirely generated during times of high mispricing. On the other hand, most of the investment factor return stems from times of low mispricing.

Looking at the conditional equity risk premium provides an additional robustness test whether the fundamentals-based mispricing model actually separates more riskier times of high mispricing and less riskier times of low mispricing. The full-sample mean of the market excess return is 0.39% p.m. but only 0.21% in times of low mispricing. As expected, the market risk premium rises to 0.64% in times of high mispricing, thus reflecting periods when stock prices are less linked to accounting fundamentals and are generally perceived more riskier by investors. Our findings in this section are in line with Baltussen et al. (2021) who document that global factor risk premiums are generally weaker in low sentiment states while unrelated with macroeconomic risks.[28] Overall, the value residual is not only able to identify mispricing across firms around the world, but is also able to identify different market-wide mispricing regimes over time.

3.5.5 Arbitrage asymmetry

As with any investment strategy that aims to exploit mispricing, arbitrageurs would correct the initial price distortion. What prevents them from these particular investment opportunities is firm specific risk such as idiosyncratic volatility (*IVOL*) (see e.g., Pontiff (2006); Stambaugh et al. (2015)).[29] To supplement the mispricing-based explanation for *VRES*, we test the empirical implication of arbitrage risk, i.e., if the according long-short return is higher among high *IVOL* stocks which are riskier for arbitrageurs and lower among low *IVOL* firms.

We calculate value-weighted returns for monthly updated portfolios which are the intersection of independent 5×5 sorts on *VRES* and *IVOL*. The first zero-cost portfolio return of interest is the difference of low- and high *VRES* quintile portfolios among stocks that all belong to the lowest *IVOL* quintile portfolio in that month. Having low idiosyncratic risk, these stocks are less susceptible to mispricing that is not eliminated by arbitrageurs. We indeed observe a value-weighted *VRES* premium of 0.32% in the same magnitude as in our portfolio sorts in Panel A of Table 3.6, so there seems to be no additional return premium besides our value residual. On the other hand, high *IVOL* stocks are more prone to mispricing, especially overpricing. For that reason, the second return of interest is the according *VRES* hedge return among stocks in the highest *IVOL* quintile portfolios. Its average return is 0.71% per month, so more than twice as high than the return among low *IVOL* portfolios.

As a further robustness test, we include *IVOL* and an interaction term between *VRES* and a dummy variable being one if a stock belongs to the highest *IVOL* quintile portfolio in that month in the cross-sectional regression shown in Spec. (1) in 3.5. As already documented in Spec. (7) to (9), controlling for *IVOL* still results in a highly significant coefficient for *VRES*. More interestingly, the estimated slope on the interaction term is a highly significantly -0.04 (t-statistic: -3.13), supporting our findings in this section that

[28]Notably, our aggregated mispricing measure is highly correlated (-0.54) with the sentiment index proposed in Baker and Wurgler (2006). The negative sign is because times of high market-wide mispricing are identified by low net of 10-year rolling average adj. R^2 values.

[29]In line with previous studies, *IVOL* in this section is calculated relative to the Fama-French three-factor model (Fama and French (1992); Fama and French (1993)) using developed ex-U.S. factor returns and we require at least 15 observations of daily stock returns over the previous month.

the mispricing related return premium is much higher among high *IVOL* stocks. Finally, most of this higher return premium stems from the poorer performance of the according short leg which comprises overpriced stocks being even more overpriced, as arbitrageurs are more reluctant to short sell them because of their higher idiosyncratic risk.

3.5.6 Transaction costs

Arbitrage capital is only able to correct potential mispricing if the trading opportunity is profitable net of transaction costs. To measure trading costs at the stock level, we apply the bid-ask spread estimator proposed in Abdi and Ranaldo (2017) that jointly considers daily high, low, and close prices to estimate the spread, i.e., $spread_t = 2\sqrt{\mathbb{E}\left[(c_t - \eta_t)(c_t - \eta_{t+1})\right]}$. The daily close log-price is denoted as c_t and η_t is the daily mid-range (average of daily high and low log-prices). The monthly estimator *spread* is the average of daily estimates where at least 15 observations within that month are required. We are able to calculate *spread* for 89.54% of our entire firm-month observations. Bid-ask spreads when direct estimates are unavailable are replaced with the estimated spread of the nearest match for which a direct spread in that month is available. As in Novy-Marx and Velikov (2016), the nearest match between two firms i and j in month t is defined by the shortest Euclidean distance in *SIZE* and *IVOL* rank space, i.e., $\sqrt{[\text{rank}(SIZE)_{i,t} - \text{rank}(SIZE)_{j,t}]^2 + [\text{rank}(IVOL)_{i,t} - \text{rank}(IVOL)_{j,t}]^2}$. At the portfolio level, transaction costs TC are estimated following Detzel et al. (2023) as $TC_t = \sum_{i=1}^{N_t} spread_{i,t} \cdot |w_{i,t} - w_{i,t-1} \cdot (1 + r_{i,t})|$, with $w_{i,t}$ as the weight of stock i in month t after rebalancing and N as the total number of stocks in the portfolio. Finally, the net-of-cost return of our portfolio strategy is defined as $r_t^{net} = r_t^{gross} - TC_t$.

Our previous results confirm that the easily exploitable long leg, holding undervalued, low-*VRES* firms, contributes the returns to the strategy. This is in line with U.S. findings, where 80% of the *VRES* strategy returns' are generated by the long leg.[30] The high value-weighted (gross) return of 0.87% per month (see Panel A of Table 3.6) for these firms makes it possible for investors to implement a long-only approach. Transaction costs lower gross returns for this global, long-only *VRES* portfolio strategy by just 8 bps, although we do not consider any cost mitigation techniques discussed in Novy-Marx and Velikov (2016). This is mainly because associated portfolio sorts reveal that 65% (50%) of stocks in the lowest (highest) quintile portfolio remain in that specific portfolio over the next four quarters, thus lowering transaction costs by low turnover. Further, the long-only

[30] This is somewhat surprising given the fact that Stambaugh et al. (2012)) shows that mispricing related profits are typically generated by the short leg.

approach avoids high shorting costs which typically lower the profitability of long-short arbitrage strategies by almost 40% without consideration of trade-related transaction costs, such as bid-ask spreads and brokerage fees (see Kim and Lee (2023)). The resulting after-cost return of 0.79% per month is highly significant with a t-statistic of 2.90.[31] Even when financed by the risk-free interest rate, this long-only investment yields an excess return of 6.96% per year net of transaction costs and outperforms the global (gross) market return by a highly significant 0.19% per month. The according alpha of 0.24% per month with respect to the Fama/French five-factor model incl. momentum confirms our main hypothesis, that low-*VRES* firms are indeed mispriced and offer higher, abnormal returns, not related with common risk factors.

3.5.7 Robustness of the fundamentals-based valuation model extension

Nichols et al. (2017) seminal framework relies on annual cross-sectional regressions linking share prices with a lag of three months to firm fundamentals, while this study conducts quarterly cross-sectional regressions as stated in Eq. (3.3). To assess the robustness of our modified framework, this section presents results using a more conservative and common lag period of six months. Additionally, we explore whether the quarterly frequency of the regressions is pivotal to our main findings and whether the need to dissect DIV^{cs} as suggested in Section 3.3.2 remains valid at the standard annual run frequency.

Table 3.9 complements our findings in Table 3.2 and shows coefficient estimates for the fundamentals-based valuation model using different regression specifications. Spec. (1) is directly related to Walkshäusl (2021) using EMU firms having their fiscal year-end on December 31 and applying annual cross-sectional regressions with financial data as of end June of the subsequent year. Our results support the main findings of Walkshäusl (2021) for the prolonged sample period and show that all accounting variables except *OIGR* have significant explanatory power for the absolute share price. We confirm the significant DIV^{cs} estimate indicating a positive dividend-price relation among EMU located firms.

[31] At the regional level, the after-cost long-only return is 0.90% in Europe and Asia-Pacific, and 0.52% in Japan.

Table 3.9: Cross-sectional regressions (1990-2021, Europe) for the fundamentals-based valuation model in Nichols et al. (2017).

This table shows time-series averages for the coefficients from firm-level cross-sectional regressions of the share price on accounting fundamentals. BV is the book value of equity, IB is income before extraordinary items, DIV^{cs} is book value of equity in year y_{t-1} minus book value of equity in year y_t plus net income in year y_t. $OIGR$ is the annual (absolute) change of operating income. All of the aforementioned variables are divided by shares outstanding. NEG is a binary indicator variable that is equal to one if IB is negative and zero otherwise. All specifications include industry dummies (not tabulated) based on Industry Classification Benchmark (ICB) categories. *Region* indicates the countries included in the sample and *FY-end* reports which firms based on the date of their fiscal year-ends are included. *Time* indicates if annual regressions or quarterly regressions are conducted, using observations of the whole calendar year or only for firms having their fiscal year-end within the last calendar quarter. For quarterly regressions, we require a minimum of at least 30 observations. *Country dummies* indicates whether country dummies are included. To avoid a look-ahead bias, our dependent variable (share price) is measured using a time lag denoted in months following the according fiscal year-end and reported in *Acc. lag*. *Month dummies* indicates if dummy variables for the month of the fiscal year-end are included. N denotes the average number of firms. The R^2 values are adjusted for degrees of freedom. Newey and West (1987) corrected t-statistics are given in parenthesis and */**/*** indicate significance at the 10%/5%/1% level. The sample period is from June 1990 to June 2021.

Specification	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
BV	0.84***	0.83***	0.83***	0.83***	0.79***	0.86***	0.84***	0.70***	0.67***	0.77***
	(15.09)	(12.96)	(16.23)	(17.14)	(19.15)	(14.36)	(13.30)	(15.26)	(11.36)	(18.21)
IB	6.79***	6.42***	6.72***	7.17***	7.36***	6.64***	6.72***	8.87***	9.60***	6.68***
	(12.73)	(12.06)	(13.20)	(12.82)	(14.83)	(11.42)	(10.02)	(13.28)	(11.09)	(14.47)
NEG	−1.57**	−1.80*	−2.05***	−2.29***	−2.32***	−3.12***	−6.09***	−1.99***	−1.98***	−2.03***
	(−2.15)	(−1.96)	(−3.38)	(−5.12)	(−6.36)	(−3.09)	(−3.17)	(−4.01)	(−3.06)	(−4.59)
$NEG \times IB$	−11.40***	−11.16***	−11.08***	−12.66***	−13.22***	−10.93***	−10.54***	−14.99***	−15.97***	−12.07***
	(−14.98)	(−14.03)	(−15.08)	(−13.73)	(−16.57)	(−7.97)	(−8.14)	(−12.85)	(−10.49)	(−15.36)
DIV^{cs}	1.50***	1.47***	1.47***	1.29***	1.16***	2.26***	2.11***	0.43	0.19	1.15***
	(5.44)	(5.35)	(5.53)	(4.71)	(4.39)	(5.19)	(4.64)	(1.12)	(0.38)	(4.29)
$OIGR$	0.38	0.20	0.37	0.50*	0.41	−0.25	−0.39	−0.32	−0.54	0.33
	(1.30)	(0.67)	(1.29)	(1.73)	(1.47)	(−0.59)	(−0.93)	(−0.82)	(−1.05)	(1.09)
Adj.R^2	0.77	0.78	0.78	0.79	0.78	0.84	0.84	0.86	0.89	0.79
N	1,195	1,185	1,195	2,174	2,946	363	363	731	231	2,228
Region	EMU	EMU	EMU	Europe	Europe	EMU	EMU	Europe	Europe	Europe
Time	Annual	Annual	Annual	Annual	Annual	Quarterly	Quarterly	Quarterly	Quarterly	Quarterly
Country dummies	No	No	Yes	No	No	No	Yes	Yes	Yes	Yes
Month dummies	No	No	No	No	No	No	Yes	Yes	Yes	Yes
FY-end	31.12.	31.12.	31.12.	31.12	All	All	All	All	Q1-Q3	Q4
Acc. lag	6	3	6	6	6	6	3	3	3	3

In contrast, we observe a significant discount on the share price for firms with negative income similar to U.S. firms. On average, the valuation penalty is € 1.57 and if income is negative, an additional loss of one euro per share decreases its price by € 4.61 (formally, € 6.79 + (- € 11.40)). Our main findings are robust against changes (1) in the accounting lag using share prices at the end of March instead of June of the subsequent year, (2) in including additional country dummies, (3) in including non-EMU firms, and (4) in including dummies for the month of a firms' fiscal year-end. Spec. (6) and (7) indicate that a quarterly regression approach provides a more detailed analysis because the time-series average of the cross-sectional adj. R^2 increases to over 80% and exceeds the amount of explainable price variation from prior studies and previous regression specifications.

Most of our non-EMU, European located firms are from the United Kingdom (UK), accounting for 30% of our total European market capitalization. In the UK, only 43% of all firms choose December 31 as their fiscal year-end and almost one out of five prefer the end of March. We try to capture the difference between UK and EMU-firms to some extent in Spec. (9) and (10) by separating all firms by the calendar quarter of their fiscal year-end. Spec. (9) includes most of UK firms and we see both economically and statistically important differences for our average DIV^{cs} estimate compared with Spec. (10). While this coefficient is within the range of previous estimates for all European firms having their fiscal year end in the fourth quarter, it is negligible for European firms having their fiscal year-end within the first three calendar quarters (which are mostly located in the UK). This raises important questions for the suitability of applying DIV^{cs} in an international sample because both specifications already include country dummies with the intent to capture possible between-country effects.

3.6 Conclusion

In this paper, we extend the fundamentals-based valuation model proposed in Nichols et al. (2017) that links share prices to accounting information to be applicable among international, non-U.S. equity markets. Our extended version of the model is able to explain 81% of share price variation among firms in global, developed countries during the period 1990-2021. We provide strong supportive evidence that firms denoted as undervalued by the model, i.e., their share price is below its fundamental value as indicated by low values of the model's estimated value residual *VRES*, significantly outperform overvalued firms. The size-adjusted return difference is a highly significant 0.56% p.m. in line with previous findings by Nichols et al. (2017) for the U.S. and Walkshäusl (2021) for the

European Monetary Union. Although undervalued firms tend to be small, value firms, none of common risk factors including size, value, profitability, or investment is able to explain the according return difference, even after controlling for further return predictors e.g., momentum, accruals, stock issuance, idiosyncratic volatility, skewness, MAX, or mispricing related controls proposed by Bradshaw et al. (2006) ($\Delta XFIN$) or Piotroski (2000) ($FSCORE$).

Our empirical findings are consistent with a portfolio strategy that exploits actual mispricing rather than gaining high returns because of exposure to risk:

First, we find that global *VRES* related strategy returns are on average 43 bps lower after publication of the original study in May 2017. This post publication return decline is expected if academic research draws investor attention to arbitrage opportunities (Hanson and Sunderam (2014); McLean and Pontiff (2016)). Because of impediments to short selling, overpricing is more difficult to eliminate compared to the very easily exploitable correction of undervaluation (which is in fact just a long position, see Stambaugh et al. (2012)). That implication is strongly supported by our finding that the return drop is entirely rooted in the long portfolio leg comprising undervalued firms and being less profitable by 39 bps after May 2017.

Second, stocks that transition out of the low *VRES* portfolio over time generate average size-adj. returns of 5.07% over the next year. Stocks that remain in this extreme portfolio on the other hand only earn -0.79% for the subsequent year. Similarly, stocks that leave (remain in) the highest *VRES* quintile experience a cumulative size-adj. return of -5.97% (3.90%) over the year following the portfolio formation month. This makes a risk-based explanation less likely because the risk implied with buying these extreme portfolio stocks would either reverse, resp. decay, or change very quickly over the period of only three months until portfolios are updated (Fama and French (2007); Jegadeesh and Titman (2001)).

Third, investment strategies that are - at least partially - associated with the exploitation of mispricing are linked with market-wide investor sentiment. In line with Stambaugh et al. (2012), the short leg of the zero-cost portfolio strategy which holds overvalued stocks is significantly negatively related with investor sentiment and shorting it is more profitable following periods of high sentiment. The global *VRES* investment strategy earns on average a highly significant (value-weighted) return of 0.41% in times of high sentiment but only a weakly significant return of 0.17% in times of low sentiment.

Finally, high idiosyncratic volatility ($IVOL$) represents risk for arbitrageurs and prevents them from particular investment opportunities that correct potential mispricing (Stambaugh et al. (2015)). We indeed observe that the *VRES* hedge return among stocks in the highest *IVOL* quintile portfolio is more than twice as high in magnitude (0.71%) than among low *IVOL* stocks (0.32%).

Our analysis reveals that investors who buy undervalued stocks and sell overvalued stocks according to the value residual proposed in Nichols et al. (2017) actually exploit mispricing opportunities among global, non-U.S. capital markets and the according size-adjusted return difference of 0.56% p.m. is not a reward for facing risk exposure.

Chapter 4

The Relevance of Risk, Mispricing, and Optionality in Momentum Returns

This research is the result of a single-author project. The paper is currently under review in *The Quarterly Review of Economics and Finance*. The journal ranking is *B* according to the VHB Publication Media Rating 2024. The paper was accepted and presented at the Doctoral Workshop of the 30th Annual Meeting of the German Finance Association (DGF) 2024 in Aachen.

Abstract

We develop a novel approach for decomposing returns of 28 equity momentum strategies into a risk-, mispricing-, and option-component. Average UMD returns of 0.66% p.m. contain: (i) an insignificant 0.29% risk-component, (ii) 0.50% mispricing-component, and (iii) -0.13% option-component. The risk-component is related with market volatility and innovations in the term spread, whereas the mispricing-component covaries with illiquidity. While standard factor models capture the risk-component of past losers across all size segments of 5×5 size-momentum portfolios, intercepts for winners are larger in magnitude compared to composed momentum returns.

Keywords: Asset pricing · Momentum · Market anomalies · Market efficiency

JEL classification: G11 · G12 · G40.

4.1 Introduction and literature review

We build upon the insight that momentum strategies are punctuated with occasional crashes (Daniel and Moskowitz (2016)) and develop a novel approach to decompose momentum returns into a risk-, mispricing-, and option-component. The mispricing-component is estimated directly in form of exposure to related factors proposed in Stambaugh and Yuan (2017), Asness et al. (2019), and Daniel et al. (2020a), whereas the residual component left unexplained by these mispricing factors is ascribed to risk, resp. optionality. A major advantage of this procedure is that we can avoid to make assumptions which of the hundreds of proposed risk factors documented in the literature (see Harvey and Liu (2016) or Hou et al. (2020)) *actually* capture exposure to risk (see Birru et al. (2023)).[1]

Rational and behavioral asset pricing models further imply a tug of war in expected returns for speculative stocks because according characteristics (e.g., small, young, lottery-like, or close to distress) overlap with stocks perceived as being risky (Baker and Wurgler (2007); Birru (2018)). Rational models predict high average returns as compensation for bearing risk while behavioral models expect low average returns as systematic mispricing predominately takes the form of overpricing in consequence of short-sale constraints and limits to arbitrage (Miller (1977); Stambaugh et al. (2012)).[2] In this paper, we analyze the momentum effect amidst this ambivalent return relation.

The main contribution of Stambaugh and Yuan (2017) and Daniel et al. (2020a) is the demonstration that mispricing related factors are beneficial in explaining the cross-section of stock returns. Even though they provide some explicit analysis on how well their factor models capture different forms of momentum, their implicit starting position is to take the momentum effect as a puzzling anomaly for granted and then benchmark the models against momentum and others. Our approach is exactly the opposite. Based on the success of these mispricing factor models, our decomposition approach begins with applying them to first reduce the noise of common mispricing in momentum returns.

[1] Hou et al. (2015) carefully acknowledge that their model *"[. . .] is silent about the debate between rational asset pricing and mispricing"* (p. 684). MacKinlay (1995) emphasizes that without a specific theory identifying risk factors, the cross-section of expected returns is always explained by some multifactor model even if non-risk based causes like market frictions and irrational investors prevail. Widely used empirical factor models are, however, often only weakly motivated by very general theories like the dividend discount model or the production-based model of Cochrane (1991).

[2] Daniel et al. (2001), Barberis and Shleifer (2003), and Baker and Wurgler (2006) provide evidence that mispricing has a common component across stocks related with time-varying market-wide investor sentiment mostly affecting speculative stocks.

Next, using these mispricing-cleaned returns, we are the first to numerically quantify the importance of mispricing relative to the remaining risk and option components in momentum returns, without ex-ante specification of particular risk factors. This differs conceptually from recent studies that also try to decompose the momentum effect. They either analyze the strategies' returns conditional on state variables (e.g., Barroso and Wang (2025)), or decompose momentum betas that represent exposure to candidate variables (e.g., Guo et al. (2022)). In either cases, state or candidate variables have to be selected a priori. Our approach, however, does not rely on prior assumptions about which of the hundreds of proposed risk factors *really* capture exposure to risk, or which variables are *actual* state variables momentum depends on.

We consider the U.S. market for the period July 1972 to December 2016 and find that the average UMD momentum return of 0.66% p.m. consists of a 0.29% risk-component (t-statistic: 1.94), 0.50% mispricing-component (t-statistic: 6.35), and -0.13% option-component (t-statistic: -3.57). Momentum is among the most studied market anomaly in asset pricing which motivated research to propose multiple versions e.g., industry or volatility-scaled momentum. For that reason, we decompose a total of 28 U.S.-equity momentum versions compiled in a comprehensive list by Chen and Zimmermann (2022) and published in top-tier journals and find that the risk-component (mispricing-component) accounts on average for 56.66% (29.63%) of returns among all 28 momentum versions.

Figure 4.1 presents return components for a selection of most common momentum versions and shows a huge dispersion in their relative importance. Most of average returns in value-momentum proposed by Novy-Marx (2013), decile 10-1, or standard (UMD) momentum is reflected in the mispricing-component. On the other hand, risk is the predominate component in volatility-scaling momentum strategies and in momentum among anomaly hedge portfolios (factor momentum). We provide evidence that the superior performance of volatility-scaling momentum strategies is explained by a risk premium for bearing aggregated market volatility risk. We next analyze how each return component covaries with macroeconomic risk, investor sentiment, and the availability of arbitrage capital. First, the mispricing-component is strongly related with sentiment, as expected, and its long leg (weakly) tends to generate higher returns when arbitrage capital is scarce. Second, we reveal that the negative relation between momentum returns and the default spread between U.S. corporate bonds as documented in Chordia and Shivakumar (2002) is almost entirely rooted in the short leg of the risk-component. Not isolating the risk-component of momentum returns from sentiment induced effects of common mispricing could potentially explain the puzzling finding in Griffin et al. (2003) that their unconditional tests fail to

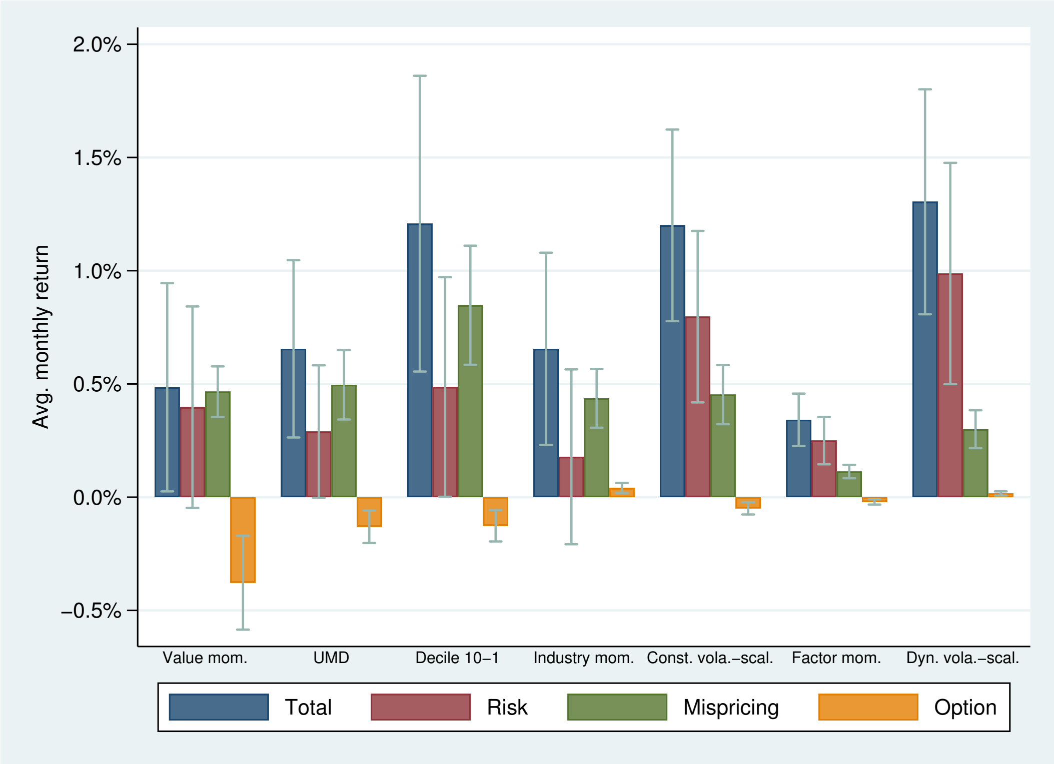

Fig. 4.1. Decomposing selected U.S.-equity momentum strategy returns into a risk-, mispricing-, and option-component. The whiskers correspond to a 95% confidence interval according to Newey and West (1987) robust standard errors. The sample period is July 1972 to December 2016.

find evidence of a relation between macroeconomic variables and (composed) momentum profits. Third, while neither leg of momentum covaries with investor sentiment, the short leg of the risk-component is strongly positive related after controlling for macroeconomic risk and even common mispricing. Economically, if one-month lagged sentiment decreases by one standard deviation (i.e., investors became pessimistic about the state of the market), monthly short leg returns in the risk-component lose an additional 0.64%, which is close to the magnitude of the overall momentum return of 0.66%. This is unexpected, because the separate option-component already captures momentum crashes which occur following market declines (i.e., when sentiment is low) and contemporaneous market rebounds.

Next, we examine asset pricing implications for decomposed momentum returns as *"momentum is a hard sell for a world of rational pricing [...]"* (Fama and French (2020), p. 1894). The Fama/French models (Fama and French (1992), Fama and French (1993), and Fama and French (2015)) or the q-factor model (Hou et al. (2015)) conclude that according factors aim to capture exposure to latent priced risk factors. Are these models able to fully describe at least the risk-component of 25 size-momentum portfolios? The answer is no. They successfully capture returns of past losers across all size segments, but average absolute intercepts for past winners are larger in magnitude and estimated more precisely compared to composed momentum returns. This finding even holds if we augment the aforementioned models with the time-series return of the risk-component.

We further analyze how momentum returns are related with liquidity and market volatility risk. In contrast to Avramov et al. (2016) who analyze momentum since 1928, we find for our shorter period 1972-2016 that momentum returns are not larger in liquid market states. The long leg of momentum generates even higher returns when aggregated market illiquidity is high (the same conclusion holds for funding illiquidity). This relation is, however, fully subsumed by the mispricing-component of momentum. This finding supports the view in Huber (2022) that strategies to exploit mispricing are difficult to implement in illiquid markets, improving the performance of the momentum strategy. Our results demonstrate that the relation between aggregated market volatility and momentum as proposed in Avramov et al. (2016) is entirely subsumed by the risk-component of momentum. This explains the superior performance of volatility-scaling momentum strategies as described in Barroso and Santa-Clara (2015), Daniel and Moskowitz (2016), and Wang and Yan (2021), which is the outcome of an effective risk management unrelated with mispricing.

Next, we test whether the risk-component of UMD momentum relates to deeper economic models in the meaning of the Intertemporal Capital Asset Pricing Model proposed by Merton (1973). We examine an empirical implementation of this model in which the factors are innovations in state variables that forecast future investment opportunities. By choosing macroeconomic variables that have forecasting power for future investment opportunities, we show that the risk-component of UMD is positively related with innovation in term spread risk which have important pricing implications. Using a machine learning approach to identify relevant economic sources for the risk-component, we find that value strategy returns are key determinants among a total of 169 macroeconomic and financial variables which serve as a channel to transmit term spread risk into UMDs' risk-component.

We assess the robustness of our results in numerous ways. First, all three return components are also important drivers of momentum returns in international, non-U.S. markets. Second, we consider momentum returns net of transaction costs (and reduced portfolio turnover to mimic real world strategy implementations) and document that the mispricing-component remains an important and well estimated driver of momentum even among the subsample of large firms. The risk-component, however, is only prominent among microcaps. Third, we reexamine our decomposition approach using daily returns. In line with Birru (2018) and Stambaugh et al. (2012), the short leg of the mispricing-component generates high returns on Friday, when sentiment and mood is at its peak during the week. The resulting long-short performance of the mispricing-component is poor on Friday (-1 bp, t-statistic: -2.71), resp. good on Monday (5 bps, t-statistic: 9,81), providing further evidence that our decomposition approach actually captures common mispricing effects. Next, van Binsbergen et al. (2023) enables us to construct a time-series of momentum returns based on rational expectations assuming the CAPM. In line with their results, our decomposition approach concludes that the long momentum leg is accurately priced at the time of portfolio formation whereas the short leg is undervalued. Lastly, simulation evidence regarding our proposed method shows that it has strong statistical power.

A huge variety of theories, both risk and behavioral models, have been proposed to explain momentum. Our paper contributes to merge these two distinct pillars of existing literature, as our findings imply that momentum returns are essentially driven by both, mispricing and risk.[3]

Risk-based explanations are presented in Conrad and Kaul (1998) (differences in stocks' expected returns), Grundy and Martin (2001) (momentum is entirely driven by residual returns and not by systematic risk components of returns), Chordia and Shivakumar (2002) (momentum is captured by conditional expected returns predicted by macroeconomic variables), Johnson (2002) (growth rate risk), Lewellen (2002) (dispersion in expected returns that is persistent over time), Zhang (2006) (information uncertainty about growth options), Sagi and Seasholes (2007) (growth options), Liu and Zhang (2008) (growth rate of industrial production), Stivers and Sun (2010) (momentum is pro-cyclical and depends on the market state), Vayanos and Woolley (2013) (flows between investment funds), Andrei and Cujean (2017) (decentralized exchange of information), Kelly et al. (2021b) (momentum characteristics are predictable by exposures to common risk factors), and Gormsen and Jensen (2024) (conditional market risk).

Behavioral explanations for momentum are proposed in Chan et al. (1996) (underreaction),

[3]In case of decile portfolio sorts on momentum, the related long-short strategy generates an average return of 1.21% p.m., of which 0.49% are attributed to the risk-component and 0.85% to the mispricing-component. The remaining -0.13% account for the option-like behavior documented in Daniel and Moskowitz (2016).

Barberis et al. (1998) (both under- and over-reaction), Daniel et al. (1998) (overconfidence), Hong and Stein (1999) (both under- and over-reaction), George and Hwang (2004) (anchoring bias), Grinblatt and Han (2005) (disposition effect), Chui et al. (2010) (self-attribution bias and overconfidence), Da et al. (2014) (inattentiveness to continuously arriving information), Hillert et al. (2014) (media coverage and overreaction), Barberis et al. (2021) (prospect theory), Hung et al. (2022) (limited attention), Huang (2022) (time-varying investor biases), Lou and Polk (2022) (investor crowding), Frey (2024) (analysts' earnings forecast), and Goyal et al. (2025) (overconfidence under the proviso that investors are more confident in rising, low-volatility markets).

The paper is organized as follows: Section 4.2 describes the data. Section 4.3 describes our decomposition approach. Section 4.4 shows our main results and investigates asset pricing implications as well as potential economic mechanisms for each momentum component. Section 4.5 provides further robustness tests and Section 4.6 concludes.

4.2 Data and summary statistics

We obtain monthly value-weighted returns for momentum sorted (decile) portfolios and returns for the Fama/French five-factor model from Kenneth French's website. The five-factor model of Fama and French (2015) (FF5 thereafter) consists of the market factor in excess of the risk-free rate $MKTRF$ and long-short portfolios to capture common covariance among stock returns related with firm size (SMB), book-to-market ratio (HML), operating profitability (RMW), and investment behavior (CMA). Built on the neoclassical q-theory of investment, Hou et al. (2015) propose a four-factor model (HXZ thereafter) consisting of a market factor (identical to $MKTRF$), size factor ME, investment factor I/A, and a profitability factor ROE. Barillas and Shanken (2018) develop a Bayesian approach to compare factor model probabilities for the collection of all possible pricing models that are based on subsets of given factors. They identify a factor model (BS thereafter) combining selected factors of FF5 (market, size, and value), HXZ (investment and profitability) and the momentum factor UMD proposed in Carhart (1997).[4] The value factor used in BS is a monthly updated version of HML, denoted as $HMLm$, and proposed in Asness and Frazzini (2013). We further consider the liquidity factor LIQ documented in Pastor and Stambaugh (2003) and the betting-against-beta factor BAB (Frazzini and Pedersen (2014)) in one of our robustness tests.

[4] As explained later, we drop the UMD factor in our analysis to avoid the circular argument of explaining momentum by momentum itself.

Similarly, Daniel et al. (2020a) introduce a mispricing model (DHS thereafter) to capture covariance with common elements of mispricing. Their two factors, *FIN* and *PEAD*, aim to capture mispricing of persistent nature (*FIN*) and of transient nature (*PEAD*). The mispricing model by Stambaugh and Yuan (2017) (SY thereafter) also introduces two factors, *PERF* and *MGMT*, based on two distinct clusters of eleven prominent anomalies. The six anomaly variables used in *MGMT* represent quantities that a firms' management can directly affect and capture time-variation in long-term mispricing. The remaining five anomalies are related to performance and thus short-term mispricing effects. Finally, the quality-minus-junk factor *QMJ* proposed by Asness et al. (2019) (AS thereafter) is related to analysts' expectation errors and captures systematic mispricing.

We obtain all monthly factor returns from the according authors' websites. Our sample period is July 1972 to December 2016 which is determined by availability of the factor returns and all returns are denominated in U.S. dollars.

Panel A of Table 4.1 reports summary statistics for decile portfolios constructed each month by sorting stocks on their cumulative 12-month performance skipping the most recent month. Over the sample period from July 1972 to Dec. 2016, the average exc. return of stocks in the highest momentum decile portfolio amounts to a highly significant 1.01% per month. Except for one portfolio (decile four), we observe monotonic increasing returns among all decile portfolios. The according long-short (L-S) portfolio investing in the highest decile portfolio (winner) and short selling the lowest one (loser) earns on average 1.21% per month (t-statistic: 3.62). Our results emphasize that most of the L-S momentum return is earned by the according long leg, as the average return of -0.19% per month for the short leg is statistically not distinguishable from zero. The median excess-return of -0.25% (1.33%) for the short (long) portfolio indicates that the according mean return is upwards (downwards) biased.

To put momentum portfolio returns in perspective, Panel B and C of Table 4.1 reports summary statistics for factor returns. The average market excess premium is 0.54% per month, only half the return of the L-S momentum return. Except for *SMB*, all risk- and mispricing-factors earn both economically and statistically significant return premiums.

Table 4.1: Summary statistics for monthly momentum and factor returns.

Panel A of this table shows mean and median of monthly value-weighted returns (in %) in excess of the one-month Treasury bill rate for monthly updated decile portfolios, sorted on cumulative past 12-month performance skipping the most recent month. The last column *L-S* reports long-short returns which are the differences between extreme decile portfolio (10-1) returns. Panel B reports mean and median of monthly returns for the risk factors proposed in Fama and French (2015) (FF5) and Hou et al. (2015) (HXZ), together with the monthly updated value factor *HMLm* proposed by Asness and Frazzini (2013), the liquidity factor *LIQ* from Pastor and Stambaugh (2003), and the betting-against-beta factor *BAB* from Frazzini and Pedersen (2014). Similarly, Panel C reports mean and median of monthly returns for mispricing factors proposed in Daniel et al. (2020a) (DHS), Stambaugh and Yuan (2017) (SY), and Asness et al. (2019) (AS). *Std* is the standard deviation of monthly excess returns and */**/*** indicate significance at the 10%/5%/1% level according to Newey and West (1987) corrected standard errors using a lag of six months. The sample period is July 1972 to December 2016.

Panel A: Monthly value-weighted momentum excess returns

	Decile Portfolio										
	1 Short	2	3	4	5	6	7	8	9	10 Long	L-S
Mean	−0.19	0.34	0.45*	0.59***	0.48**	0.54***	0.58***	0.72***	0.74***	1.01***	1.21***
Median	−0.25	0.46	0.31	0.80	0.67	0.78	0.84	0.91	1.18	1.33	1.71
Std	8.41	6.43	5.52	4.91	4.57	4.59	4.43	4.48	4.85	6.21	7.31
t-stat	−0.48	1.14	1.80	2.68	2.40	2.59	3.22	3.60	3.39	3.64	3.62

Panel B: Monthly risk factor premiums

		FF5				HXZ			other		
	MKTRF	SMB	HML	RMW	CMA	ME	I/A	ROE	HMLm	LIQ	BAB
Mean	0.54***	0.20	0.41***	0.28**	0.35***	0.27**	0.41***	0.54***	0.39**	0.40**	0.91***
Median	0.86	0.10	0.32	0.26	0.21	0.21	0.36	0.65	0.18	0.30	1.13
Std	4.55	3.03	2.93	2.33	1.95	3.09	1.85	2.59	3.59	3.49	3.51
t-stat	2.58	1.55	2.66	2.43	3.62	2.07	4.62	4.88	2.19	2.47	4.59

Panel C: Monthly mispricing factor premiums

	DHS		SY		AS
	FIN	PEAD	PERF	MGMT	QMJ
Mean	0.79***	0.62***	0.63***	0.66***	0.40***
Median	0.73	0.67	0.59	0.65	0.36
Std	3.89	1.88	3.92	2.81	2.35
t-stat	4.39	7.61	3.46	4.94	3.37

4.3 Decomposing momentum returns: The baseline decomposition approach

We combine the novel return-decomposition approach presented in Birru et al. (2023) with the method in Daniel and Moskowitz (2016) to decompose momentum returns into a risk-, mispricing-, and option-component. The latter reflects the fact that the high average return in momentum strategies is punctuated with occasional crashes, i.e., extraordinary negative returns mostly in times of overall stock market panic states and high market volatility. Most of the negative returns in momentum crashes stems from the high returns of past losers that are shorted in momentum strategies: The worst month in our sample period was April 2009 (during the global financial crisis) with a loss of 45.21%. In that month, the loser decile returned 44.89% while the winner decile had a loss of -0.32%. Daniel and Moskowitz (2016) emphasize that these momentum crashes are described by an option-like behavior: In bear markets with contemporaneous market rebounds, the momentum strategy behaves as if it is effectively short a call option on the market portfolio.

The validity and robustness of our baseline decomposition approach is demonstrated in multiple ways: First, by using simulation evidence presented in Section 4.5.5. Second, to develop a full picture of our decomposition approach reflecting both the insights in Daniel and Moskowitz (2016) and Birru et al. (2023), we apply several econometric procedures to evaluate the robustness of our results with respect to changes in the baseline model. These results are presented in Section A.I of the Appendix. Third, we provide further robustness tests using principal component analysis (PCA) in Section A.II and the instrumented principal component analysis (IPCA) by Kelly et al. (2019) in Section A.III of the Appendix.

We start our decomposition analysis with the baseline model. For each of our three mispricing models (DHS, SY, AS), we regress momentum returns on the mispricing factor(s) and market indicator variables, according to the following time-series regression:

$$r^e_{i,t} = \left(\alpha_{i,j} + \alpha^B_{i,j} \cdot I_{B,t-1}\right) + \left(\beta^0_{i,j} + I_{B,t-1}\left(\beta^B_{i,j} + I_{U,t} \cdot \beta^{B,U}_{i,j}\right)\right) r^e_{m,t} + \beta_{i,j} X_{j,t} + \epsilon_{i,j,t},$$

or rearranged,

$$r^e_{i,t} = \underbrace{\beta_{i,j} X_{j,t}}_{\text{mispricing}} + \underbrace{\beta^{B,U}_{i,j} I_{B,t-1} I_{U,t} r^e_{m,t}}_{\text{option}} + \underbrace{\left(\alpha_{i,j} + \alpha^B_{i,j} \cdot I_{B,t-1}\right) + \left(\beta^0_{i,j} + I_{B,t-1}\beta^B_{i,j}\right) r^e_{m,t} + \epsilon_{i,j,t}}_{\text{risk}}, \tag{4.1}$$

where $r^e_{i,t}$ is the value-weighted return of momentum decile portfolio i in month t in excess of the risk-free rate rf_t (one-month Treasury bill rate) and $\alpha_{i,j}$, $\beta_{i,j}$, and $X_{j,t}$ are vectors of intercepts, loadings, and mispricing factors, respectively, corresponding to mispricing model j. The mispricing-component of portfolio returns is the sum of the product(s) of the estimated loading(s) and factor(s) $\hat{\beta}_{i,j}X_{j,t}$. In line with Daniel and Moskowitz (2016), the option-component is the product $\hat{\beta}^{B,U}_{i,j}I_{B,t-1}I_{U,t}r^e_{m,t}$, where $r^e_{m,t}$ is the market (CRSP value-weighted index) excess return. $I_{B,t-1}$ is an indicator variable (bear market) that equals one if the cumulative market return in the past 24 months is negative and zero otherwise. $I_{U,t}$ is an indicator variable (up market) that equals one if $r^e_{m,t}$ is positive and zero otherwise. The risk-component is the sum of the fitted values of the remaining part that includes intercept, residuals, and exposure to market risk. Finally, we average the risk-, mispricing-, and option-components across the three factor models.

There is an important point that needs to be discussed. Our main attempt is to isolate and remove the component of momentum returns related to mispricing factors. Accordingly, the first term of Equation (4.1) is straightforward and closely follows Birru et al. (2023). The decision of assigning the intercept $\alpha_{i,j}$ to the "risk-component" seems unconventional at first glance (again, this closely follows Birru et al. (2023)), because commonly, intercepts or "alphas" are not a measure of risk, but of risk-adjusted returns. To demonstrate that this is, however, conceptually right, think about the decomposition of an undisputed source of risk, the market portfolio itself. Its decomposition according to Equation (4.1) collapses into a pure risk-component in precisely the magnitude of the average market excess return (0.54%), while both the mispricing- and option-component rightly equal zero.[5] Of course, what we term the "risk-component" can still include any effects not captured by the mispricing factors. Nevertheless, any effects of common mispricing should, however, be substantially weaker among this risk-component than should be expected for returns that are not purged of systematic mispricing reflected in the factors of DHS, SY, and AS.

4.4 Main results

4.4.1 Risk-, mispricing-, and option-components in momentum returns

Average risk-, mispricing-, and option-components of value-weighted decile momentum portfolio excess returns using our decomposition approach described in Section 4.3 are reported in Table 4.2.

[5] See Section 4.5.5 for further details.

Table 4.2: Risk-, mispricing-, and option-components of momentum sorted decile portfolios.

This table reports monthly average risk-, mispricing-, and option-components (in %) of value-weighted decile momentum portfolio excess returns (first row). We regress momentum returns on mispricing factors proposed in Stambaugh and Yuan (2017), Asness et al. (2019), and Daniel et al. (2020a), interaction terms indicating distinct market states (bear- and bull-market), and the market return, as proposed in Daniel and Moskowitz (2016) and Birru et al. (2023). The according fitted values correspond to the mispricing-, and option-component, while the remainder (intercept, residuals, and market exposure) corresponds to the risk-component. Finally, we average each component across the three factor models (DHS, SY, AS). The sample period is July 1972 to December 2016. */**/*** indicate significance at the 10%/5%/1% level according to Newey and West (1987) corrected standard errors.

	Decile Portfolio											
	1 Short	2	3	4	5	6	7	8	9	10 Long	Long-Short Mean	 t-stat.
Avg. return	−0.19	0.34	0.45*	0.59***	0.48**	0.54***	0.58***	0.72***	0.74***	1.01***	1.21***	3.62
Risk	0.47	0.64**	0.52**	0.61***	0.46**	0.56**	0.40**	0.48**	0.58**	0.95***	0.49**	1.97
Mispricing	−0.77***	−0.37***	−0.18***	−0.06***	−0.00	0.07***	0.18***	0.19***	0.20***	0.07**	0.85***	6.32
Option	0.11***	0.07***	0.11***	0.04***	0.03***	−0.09***	0.01***	0.05***	−0.05***	−0.01***	−0.13***	−3.57

The major portion of the 1.21% long-short momentum return stems from the highly significant spread return of 0.85% (t-statistic: 6.32) for the mispricing-component.[6] The risk-component shows no clear pattern related with decile portfolios, but sharply increases from 0.58% (decile 9) to 0.95% for the winner portfolio. The mispricing-component in momentum decile portfolios is strictly monotonic increasing from -0.77% for decile 1 to 0.20% for decile 9 and then drops to 0.07% for the winner portfolio. This monotonic relationship among all but decile 10 portfolios is highly significant (p-value: 0.008) using the non-parametric test for monotonicity in asset returns proposed in Patton and Timmermann (2010). The mispricing-component is unevenly prevalent among decile portfolios. The corresponding 5-1 return differential is a highly significant 0.77% (t-statistic: 6.08) while the 10-6 return difference is an insignificant 0.01%. The difference in differences is a highly significant -0.76% (t-statistic: -5.09) and implies that mispricing effects in momentum returns are disproportionately stemming from past losers. Similarly, the 5-1 return difference of the risk-component is an insignificant -0.01% and a significant 0.39% (t-statistic: 2.55) for decile 10-6. However, the according difference in differences of 0.39% is insignificant.

Looking at the importance of the risk- and mispricing-component over time, we notice that neither of them clearly dominates as both of them show substantial variation. The portion of sample months where the risk-component is larger in magnitude then the mispricing-component is a merely 49.81%. However, as shown in Figure A.I in the Appendix, the 20-year moving averages of the risk-component are downloads sloping in line with overall momentum profits: The 20-year average momentum return until June 1992 is 1.65% and 0.83% are reflected in the risk-component. Until December 2016, momentum profits decline to only 0.77%, mostly because the risk-component's return plunges to 0.16%. In contrast, the magnitude of the mispricing-component remains to be rather stable over time, contributing 0.92% in June 1992 and 0.79% in December 2016 when looking at 20-year rolling windows.

We confirm the findings in Daniel and Moskowitz (2016) that the loser-portfolio has a highly significant market exposure in bear markets with contemporaneous upward market swings (see Section A.I of the Appendix). This pattern also holds for less extreme portfolios such as decile 2 and decile 3 and accounts on average for 11 bps, resp. 7 bps, as a premium for this option-like behavior. In other words, low-momentum portfolios earn a

[6]This is in line with Kang (2024) who propose a reduced-form asset pricing model and find that the momentum phenomenon is partially caused by mispricing that does not vary with macroeconomic variables.

positive premium for being long a call option on the market return, paying off in times of momentum crashes. This optionality does not apply to high momentum portfolios, so the combined L-S momentum strategy is effectively shorting the call option, accounting for an average loss of -0.13% per month.

After having established that the long-short momentum strategy using decile portfolios is mostly driven by a mispricing-component, a relatively weaker risk-component, and a non-negligible option-component, the key question is how important are these components in other momentum strategies? The momentum effect is perhaps the most studied anomaly, and top-tier finance literature has established a vast number of related strategies. To mitigate concerns of picking and choosing momentum strategies that suit our previous results, we consider a comprehensive list of thirteen established equity momentum strategies compiled by Chen and Zimmermann (2022) that are classified as clear predictors. We augment this list with fifteen other momentum strategies that are proposed in highly cited studies among top-tier finance journals, particularly in the Journal of Finance, the Journal of Financial Economics, or the Journal of Banking & Finance. To begin with, these momentum styles comprise all eight strategies analyzed in Ehsani and Linnainmaa (2022), i.e., industry-adjusted momentum (Cohen and Polk (1998)), Sharpe ratio momentum (Rachev et al. (2007)), individual factor momentum, and momentum among principal components of factor returns. Value-weighted long-short (decile 10-1) returns for value-momentum and value-momentum-profitability as proposed in Novy-Marx (2013) are provided by Serhiy Kozaks' equity anomaly data (Kozak et al. (2018) and Haddad et al. (2020)). We also include the equally combined strategy of investing in rank-weighted value and momentum spread portfolios as suggested in Asness et al. (2013). Last, we include three volatility scaling momentum strategies proposed in Barroso and Santa-Clara (2015), Daniel and Moskowitz (2016), and Wang and Yan (2021). To make the returns of the scaled and unscaled strategies comparable, we chose the annualized target volatility to match the full sample volatility of UMD returns instead of using the commonly proposed fixed target of 12%.[7] Parameters for dynamic-scaled momentum in Daniel and Moskowitz (2016) are estimated out-of-sample following the approach in Hanauer and Windmüller (2023) to mitigate concerns of a look-ahead bias.[8]

[7]Because UMD is a zero-investment strategy, we can scale it without constraints.

[8]We thank all authors for sharing their data and Pedro Barroso for sharing his Matlab code.

Table 4.3: Risk-, mispricing-, and option-components of equity momentum strategies.

This table reports monthly average risk-, mispricing-, and option-components (in %) of 28 equity momentum strategies ranked in descending order by their average return. All momentum strategies are either classified as clear predictors for the cross-section of stock returns in a comprehensive list compiled by Chen and Zimmermann (2022) or are proposed in highly cited studies among top-tier finance journals, particularly in the Journal of Finance, the Journal of Financial Economics, or the Journal of Banking & Finance. We regress (using our baseline approach described in the main text) monthly value-weighted momentum returns on mispricing factors proposed in Stambaugh and Yuan (2017), Asness et al. (2019), and Daniel et al. (2020a), interaction terms indicating distinct market states (bear- and bull-market), and the market return as proposed in Daniel and Moskowitz (2016). Following Birru et al. (2023), the according fitted values correspond to the mispricing- and option-component, while the remainder (intercept, residuals, and market exposure) corresponds to the risk-component. Finally, we average each component across the three factor models. The sample period for all strategies ends in December 2016. */**/*** indicate significance at the 10%/5%/1% level according to Newey and West (1987) corrected standard errors.

			Monthly returns				t-statistic			
Momentum Definition	Reference	Start	Avg.	Risk	Misp.	Opt.	Avg.	Risk	Misp.	Opt.
Firm age	Zhang (2006)	12/1973	1.63***	1.37***	0.50***	−0.23***	6.14	5.97	6.33	−3.58
High volume stock mom.	Lee and Swaminathan (2000)	07/1972	1.33***	0.97***	0.82***	−0.47***	3.68	3.48	6.18	−3.57
Dyn. vola.-scaled	Daniel and Moskowitz (2016)	07/1972	1.30***	0.99***	0.30***	0.02***	5.15	3.96	7.03	3.57
Decile 10-1	Jegadeesh and Titman (1993)	07/1972	1.21***	0.49**	0.85***	−0.13***	3.62	1.97	6.32	−3.57
Const. vola.-scaled	Barroso and Santa-Clara (2015)	07/1972	1.20***	0.80***	0.45***	−0.05***	5.56	4.13	6.80	−3.57
Const. semi-vola.-scaled	Wang and Yan (2021)	07/1972	1.13***	0.74***	0.44***	−0.05***	5.18	3.74	6.87	−3.57
Junk stock	Avramov et al. (2007)	12/1978	0.99***	0.81***	0.48***	−0.29***	3.73	3.80	5.39	−3.06
FF3 residual	Blitz et al. (2011)	07/1972	0.98***	0.76***	0.34***	−0.13***	6.29	5.81	6.75	−3.57
Off season mom.	Heston and Sadka (2008)	07/1972	0.91***	1.07***	−0.19***	0.04***	3.53	4.55	−3.19	3.57
Principal customer	Cohen and Frazzini (2008)	07/1977	0.84***	0.58**	0.09***	0.17***	3.25	2.17	5.92	3.11
Return seasonality	Heston and Sadka (2008)	07/1972	0.76***	0.71***	−0.00	0.05***	5.34	4.96	−0.29	3.57
Long-term rev./mom.	Chan and Kot (2006)	07/1972	0.74***	0.48**	0.52***	−0.27***	3.12	2.16	7.63	−3.57
Value-mom.-profitability	Novy-Marx (2013)	07/1972	0.69***	0.59***	0.44***	−0.34***	3.14	2.78	7.50	−3.57
UMD (standard mom.)	Jegadeesh and Titman (1993)	07/1972	0.66***	0.29*	0.50***	−0.13***	3.28	1.94	6.35	−3.57
Industry	Moskowitz and Grinblatt (1999)	07/1972	0.65***	0.18	0.44***	0.04***	3.02	0.90	6.59	3.57
Customer industry	Menzly and Ozbas (2010)	02/1986	0.63***	0.71***	0.05**	−0.14***	3.06	3.64	2.27	−2.93
Sharpe ratio	Rachev et al. (2007)	07/1972	0.60***	0.26**	0.40***	−0.06***	3.71	1.97	6.51	−3.57
Intermediate	Novy-Marx (2012)	07/1972	0.55***	0.25	0.16***	0.14***	3.37	1.60	4.64	3.57
Value-mom.	Novy-Marx (2013)	07/1972	0.49**	0.40*	0.47***	−0.38***	2.07	1.75	8.16	−3.57
Supplier industry	Menzly and Ozbas (2010)	02/1986	0.48**	0.32	0.04	0.12***	2.27	1.58	1.02	2.93
50/50 value/mom.	Asness et al. (2013)	07/1972	0.44***	0.29***	0.23***	−0.09***	5.22	4.30	6.28	−3.57
Industry-adj.	Cohen and Polk (1998)	07/1972	0.37***	0.23**	0.23***	−0.08***	3.04	2.43	5.78	−3.57
Factor mom.	Ehsani and Linnainmaa (2022)	07/1972	0.34***	0.25***	0.11***	−0.02***	5.79	4.68	7.53	−3.57
PC factor mom. 1-10	Ehsani and Linnainmaa (2022)	01/1973	0.20***	0.15***	0.04***	0.00***	6.01	5.01	7.29	3.58
PC factor mom. 11-20	Ehsani and Linnainmaa (2022)	01/1973	0.13***	0.12***	0.03***	−0.01***	4.72	4.32	7.58	−3.58
PC factor mom. 21-30	Ehsani and Linnainmaa (2022)	01/1973	0.11***	0.06**	0.02***	0.03***	4.54	2.48	5.58	3.58
PC factor mom. 31-40	Ehsani and Linnainmaa (2022)	01/1973	0.11***	0.09***	0.04***	−0.02***	4.51	3.80	7.28	−3.58
PC factor mom. 41-47	Ehsani and Linnainmaa (2022)	01/1973	0.09***	0.12***	0.02***	−0.04***	3.03	4.18	3.85	−3.58

Table 4.3 presents results for decomposing returns for the set of 28 U.S.-equity momentum strategies into risk-, mispricing-, and option-components using the baseline approach described in Section 4.3.[9] We observe a huge dispersion in average returns among all analyzed strategies. Factor momentum performs quite poor whereas firm age momentum (Zhang (2006)) generates a remarkably high return of 1.63%. The average UMD return (Jegadeesh and Titman (1993); Carhart (1997)) is 0.66%, nearly half of the 1.21% generated by the long-short decile portfolio strategy.

A huge portion of returns in momentum strategies controlling for the value effect (Asness et al. (2013), Novy-Marx (2013)) is attributable to the mispricing-component. Given the close ties between the value premium and time-varying risk premiums, we observe that the option-component is larger in magnitude for value-momentum strategies and similarly for junk stock momentum (Avramov et al. (2007)).[10] The highest relevance of the mispricing-component is detected in industry momentum (Moskowitz and Grinblatt (1999)), i.e., investing long (short) in stocks from three winning (losing) industries as indicated by two-digit SIC codes. The average return of 0.65% stems from a highly significant 0.44% for the mispricing-component, a highly significant 0.04% for the option-component, and an insignificant 0.18% for the risk-component. However, in 21 out of 28 strategies, the risk-component dominates the mispricing-component, especially among customer-/supplier-based momentum strategies (Cohen and Frazzini (2008); Menzly and Ozbas (2010)) and in factor momentum (Ehsani and Linnainmaa (2022)). Momentum controlling for patterns in past long-term returns as documented in Heston and Sadka (2008) is entirely dominated by the risk-component. The according off-season momentum strategy with an average return of 0.91% is the only case, where the mispricing-component is highly significantly negative (-0.19%). Both results are, however, in line with findings in Keloharju et al. (2016) who document that seasonality strategies are immensely risky because of their exposures to systematic factors. Given the statistical and economical significance of our findings in Table 4.3, we conclude that both the risk-, and mispricing-component is an important source for momentum strategies.[11]

[9]We also apply all other decomposition methods outlined in Section A.I for robustness. Our results remain unchanged.

[10]The value premium is high (low) when economic conditions are poor (good) and risk premia are high (low), see e.g., Fama and French (1995), Lettau and Ludvigson (2001), and Zhang (2005).

[11]The option-component is also an important source for momentum strategies but only rarely pays off. E.g., in UMD, its return is zero in 489 out of 534 months in our sample period. However, conditional on being non-zero, the mean return is an important -1.55%.

4.4.2 Investor sentiment and arbitrage capital

If the mispricing-component is, at least partially, associated with the exploitation of mispricing, its returns should be affected by investor sentiment as demonstrated in Stambaugh et al. (2012). To be precise, we should observe a greater short leg sensitivity because of arbitrage asymmetry which leaves more uncorrected overpricing than uncorrected underpricing.[12] The performance of mispricing-related investment strategies, however, depends on the availability of arbitrage capital. A shortage of arbitrage capital in the market could potentially weaken the performance of mispricing-related strategies as arbitrage forces may not be able to correct the existing level of market-wide mispricing. To measure the availability of arbitrage capital, we use the noise index proposed in Hu et al. (2013) who conclude that this measure proxies for market-wide liquidity and the availability of arbitrage capital.[13] The noise index is based on the aggregated level of differences between market prices of U.S. Treasury bonds and model-implied yields, thus high levels of noise indicate a shortage of arbitrage capital in financial markets. We use the monthly index constructed by Baker and Wurgler (2006) to proxy for investor sentiment.

To begin with, we separately decompose the long- and short-leg of UMD into its risk-, mispricing-, and option-component using our baseline approach described in Section 4.3.[14] The average return of UMD is 0.66%, generated by 1.34% for the long leg while the short leg yields 0.69% (see Panel A of Table 4.4). Looking at the risk-component of UMD, we see that the long- (short-) leg generates a highly significant return of 0.99% (0.70%). The average return of 0.50% for the mispricing-component is entirely gained by the short leg.

Panel B shows coefficient estimates from monthly time-series regressions of UMD returns on one-month lagged levels of investor sentiment. UMD and both its legs are unrelated with sentiment, as already documented in previous studies (e.g., Jacobs (2015) and Keloharju et al. (2016)). Slope coefficients on sentiment are uniformly negative, consistent with market-wide sentiment effects. It is worth noting that the coefficient estimate of -0.27 for the short leg of the mispricing-component is smaller in magnitude compared with the -0.38 for the UMD short leg, but estimated much more precisely. Overall, the

[12] Abhyankar et al. (2024) document that momentum returns declined more recently which coincides with a significant growth in the number of stocks with (short-) options which reduces short-sale constraints. This should, however, not affect our results because our sample period starts in July 1972 and less than one third of all stocks traded had associated options until 1996.

[13] The noise index is provided on a daily basis since January 1987 and we convert the time-series by calculating monthly averages.

[14] We only report results for the risk- and mispricing-component in this section for brevity and because the option-component only rarely generates non-zero returns in times of momentum crashes.

Table 4.4: Investor sentiment, arbitrage capital, and momentum.

This table examines time-varying effects of decomposed momentum returns. Panel A shows the mean of monthly returns (in %) for the UMD factor (Jegadeesh and Titman (1993); Carhart (1997)) and its according long- and short-leg separated into risk- and mispricing-components. Panel B provides results from regressing UMD returns on the one-month lagged sentiment index proposed in Baker and Wurgler (2006). Panel C shows results from regressing UMD returns on the one-month lagged sentiment index, one-month lagged levels of the noise index (a market-wide indicator of liquidity and arbitrage capital across markets) as proposed in Hu et al. (2013), and various measures of macroeconomic risks. The macroeconomic variables are contemporaneous GDP growth rates (from NIPA), the equity market excess return (MKTRF), and the bond factor returns TERM and DEF proposed in Fama and French (1993), which represent the term spread on U.S. government bonds (10Y minus 1Y, from FRED), and the default spread measured as the difference in yields between BAA and AAA rated U.S. corporate bonds (from FRED), respectively. Because of limited data availability for the noise index, the regression estimates in Panel C are measured over the period 02/1987 to 12/2016. To obtain decomposed UMD returns, we use our baseline approach described in the main text (Section 4.3). Newey and West (1987) corrected t-statistics are given in parenthesis and */**/*** indicate significance at the 10%/5%/1% level. The sample period is July 1972 to December 2016.

	UMD	UMD		UMD^{RISK}		UMD^{MISP}	
		Long	Short	Long	Short	Long	Short
Panel A: Average monthly returns							
	0.66***	1.34***	0.69**	0.99***	0.70***	0.03	−0.47***
	(3.28)	(5.56)	(2.33)	(4.21)	(2.93)	(1.37)	(−6.28)
Panel B: Time-series regressions on one-month lagged sentiment							
Intercept	0.66***	1.33***	0.67**	0.98***	0.70***	0.02	−0.48***
	(3.38)	(5.56)	(2.33)	(4.21)	(2.96)	(1.30)	(−6.94)
Sentiment	0.19	−0.19	−0.38	−0.25	−0.10	−0.03**	−0.27***
	(1.01)	(−0.76)	(−1.17)	(−0.99)	(−0.38)	(−2.15)	(−3.47)
Panel C: Time-series regressions on one-month lagged sentiment and controls (start: 02/1987)							
Intercept	4.28**	1.34**	−2.94**	0.34	−2.12**	0.18	−1.13***
	(2.48)	(2.29)	(−2.26)	(0.76)	(−2.30)	(1.28)	(−2.68)
Sentiment	−0.22	0.23	0.45	0.31	0.70***	−0.07*	−0.30***
	(−0.54)	(1.28)	(1.31)	(1.64)	(2.62)	(−1.85)	(−2.69)
Noise	0.23	0.18**	−0.05	0.08	−0.06	0.04**	−0.04
	(1.27)	(1.96)	(−0.39)	(1.03)	(−0.69)	(1.98)	(−0.93)
MKTRF	−0.23*	1.06***	1.29***	1.06***	1.05***	0.02***	0.22***
	(−1.79)	(24.87)	(13.89)	(28.83)	(16.99)	(2.55)	(5.87)
TERM	−0.03	0.02	0.04	0.19*	0.03	−0.01	0.09
	(−0.16)	(0.15)	(0.36)	(1.94)	(0.29)	(−0.60)	(1.33)
DEF	−4.18**	−1.45**	2.74*	−0.62	2.32**	−0.28*	0.51
	(−2.19)	(−2.11)	(1.92)	(−1.14)	(2.31)	(−1.68)	(1.08)
GDP growth	−0.65	−0.06	0.59	0.00	0.37	−0.02	0.20
	(−1.11)	(−0.28)	(1.21)	(0.00)	(1.08)	(−0.48)	(1.05)

asymmetry in sentiment effects provides evidence that any remaining effects of systematic mispricing are substantially weaker in the risk-component of UMD returns, in line with the intention of our decomposition approach.

Panel C of Table 4.4 examines time-varying sentiment effects of decomposed UMD returns after controlling for the availability of arbitrage capital and various measures of macroeconomic risk following Asness et al. (2013). The macroeconomic variables are contemporaneous GDP growth rates (from NIPA), the equity market excess return, and the bond factor returns TERM and DEF proposed in Fama and French (1993), which represent the term spread on U.S. government bonds (10Y minus 1Y, from FRED) and the default spread between BAA and AAA rated U.S. corporate bonds (from FRED), respectively. Because the noise index is generally not available prior to Jan. 1987, coefficients are estimated using the period 02/1987 to 12/2016. Similar to Asness et al. (2013), momentum is unrelated with GDP growth and TERM, while the default spread negatively covaries with UMD returns. The long leg of momentum is weakly related with noise and the positive relation is more prevalent in the long leg of the mispricing-component. This finding is as expected, because if arbitrage capital is very scarce, even the long leg is able to generate higher returns, although mispricing in general takes the form of undervaluation which is quite easy to exploit. To our surprise, we observe a highly significant positive slope coefficient estimate for sentiment in the short leg of the risk-component even after controlling for systematic mispricing. Economically, a one standard deviation increase in sentiment corresponds to an increase in monthly short leg returns of 0.64%, resp. 7.64% per year. In other words, the increase in returns alone is already close to the magnitude of the average UMD return of 0.66%. We further document that the short leg of the risk-component tends to generate high returns if the default spread increases and past sentiment was high. High returns for the short leg of momentum are, however, what Daniel and Moskowitz (2016) refer to as momentum crashes. Taken together, these results suggest that the average return of 0.29% for the risk-component of UMD is indeed a risk premium, i.e., for bearing the risk of momentum crashes.

4.4.3 Asset pricing implications

After having established that both mispricing and risk are important components of equity momentum strategy returns, the key question is what do our results tell us about potential economic mechanisms for the momentum effect in context of standard asset pricing models?

Fama and French (2020) state that *"momentum is a hard sell for a world of rational pricing [...]"* (p. 1894). This raises the question, how well do leading factor models perform to explain each of these components, especially the risk-component?[15] In particular, FF5 is motivated by the dividend-discount model, whereas HXZ relies on the q-theory investment pricing model. Both aim to capture systematic risks for which investors require compensation. For that reason, we expect these risk-based factor models to perform better in explaining the risk-component of momentum returns than the mispricing-component. Hou et al. (2015) and Hou et al. (2021) claim that their model (HXZ, resp. augmented with an expected investment growth factor EG) captures momentum through the profitability factor ROE and outperforms FF5 in explaining momentum. We can address this assertion by analyzing if HXZ actually captures the return dispersion across the risk-component of momentum returns, or if the model's superiority is the result of capturing part of the mispricing-component. A reasonable ground for HXZ to potentially capture the mispricing-component is already stated in the abstract of Hou et al. (2019): The mispricing factors MGMT and PERF proposed in Stambaugh and Yuan (2017) are close to the q-factors ROE and I/A, with correlations of 0.80 and 0.84.[16]

A first glance of the performance of these empirical factor models to capture the returns of UMD's risk-component is obtained by spanning regressions. We obtain only weekly significant intercepts if we regress the risk-component of UMD on the market factor or FF5 factors, and an insignificant intercept when using HXZ factors (see Table A.V in the Appendix for details). Thus, a central implication of our decomposition approach holds: The risk-component of UMD returns is explained to a high degree by common empirical risk-based models. This simple test, however, does not reflect that firm size vastly differs across momentum sorted portfolios, an important fact already documented by Jegadeesh and Titman (1993). The average market capitalization among the lowest decile momentum portfolio is \$376.5 mio., whereas top decile firms comprise an average market capitalization of \$1,446.1 mio. To account for these size differences, our playing field in this section are the 25 size-momentum portfolios from Kenneth French's data library which are the intersections of five portfolios formed on size (market equity) and five portfolios formed on momentum.

[15]We do not intend to step into the debate which factor model takes the first place in the overall horse race to explain anomalies or what framework is best suited to compare them, as studied, e.g., in Ahmed et al. (2018), Barillas and Shanken (2018), Fama and French (2018), Barillas et al. (2020), Feng et al. (2020), Chib et al. (2020), Bryzgalova et al. (2023), and Detzel et al. (2023). Our primary interest is to evaluate if risk-based factor models are at least able to explain the risk-component in momentum returns.

[16]In addition, Novy-Marx (2015) raises concerns that ROE incorporates the most recent earnings information from quarterly data and thus reflects momentum in firm fundamentals per construction.

We closely follow Fama and French (2016) and Fama and French (2020) and analyze how well time-series factors explain excess returns for the 25 size-momentum portfolios by reporting summary statistics for the intercepts.[17] Our model performance metrics include the GRS statistic of Gibbons et al. (1989) and its key metric $Sh^2(a) = a'\Sigma^{-1}a$, i.e., the maximum squared Sharpe ratio for the vector of intercepts a. Σ denotes the covariance matrix for the residuals. Using A and V to indicate a cross-section average and variance, we also report $A|a|$ and $A|t\,(a)|$, the averages across the 25 size-momentum portfolios of the absolute intercepts and the absolute t-statistic for the intercepts. Further, we estimate the proportion of the cross-section dispersion in average returns missed by a model in two ways: The average of squared intercepts over the cross-sectional variance of the average returns $\bar{r}$ on the 25 size-momentum portfolios $Aa^2/V\bar{r}$, and $A\lambda^2/V\bar{r}$, with $\lambda^2 \equiv a^2 - s^2(a)$. By subtracting the squared standard error $s^2(a)$ from each intercept, the latter metric accounts for noise in estimated intercepts. To estimate the proportion of the dispersion of the intercept estimates attributable to sampling error, we report $As^2(a)/A(a^2)$. As a measure for time-series regression fit, we report the average regression R^2, the average standard error of the intercepts $As(a)$, and the average of residual standard deviations $As(e)$. Low values of $Aa^2/V\bar{r}$ and $A\lambda^2/V\bar{r}$ imply a good performance for a model to describe returns among the 25 size-momentum portfolios, because they indicate that intercept dispersion is low relative to the dispersion of average portfolio returns. Similar, a model that produces the smallest $Sh^2(a)$ is superior among competing models. On the other hand, high values of $As^2(a)/A(a^2)$ are good news for a model, because this implies that much of the dispersion of the intercept estimates is due to sampling error rather than to dispersion of the true intercepts.

[17] In addition to the factors in FF5 and HXZ, we also analyze characteristic-efficient FF5 portfolios introduced in Daniel et al. (2020b) that remove potential unpriced risk from FF5 factors. We thank the authors for making these data publicly available.

Table 4.5: Explaining excess returns of 5×5 size-momentum sorted portfolios.

This table examines how well time-series factors explain excess returns for 25 portfolios formed at the end of each month using independent 5×5 sorts on size (market capitalization) and momentum by reporting summary statistics for the intercepts from time-series regressions. Panel A shows results for compound returns, while Panel B and C show results for according risk- and mispricing-components in size-momentum portfolio returns, respectively. To obtain these components, we use our baseline approach described in the main text (Section 4.3). For model comparison, we use the five-factor model (FF5) proposed in Fama and French (2015) which consists of the market factor in excess of the risk-free rate *MKTRF* and factor portfolios related with firm size (*SMB*), book-to-market ratio (*HML*), operating profitability (*RMW*), and investment behavior (*CMA*). We augment the FF5-model with a momentum factor UMD (Jegadeesh and Titman (1993), Carhart (1997)) or its risk- or mispricing-component UMD^{RISK}, resp. UMD^{MISP}. We also use the according characteristic-efficient FF5 factor returns (marked with asterisks) proposed in Daniel et al. (2020b) that remove unpriced risk from FF5 factors. Finally, we use the four-factor q-theroy model (HXZ) proposed in Hou et al. (2015) consisting of a market factor (*MKTRF*), size factor *ME*, investment factor *I/A*, and a profitability factor *ROE*, resp. its augmented version with an expected (investment) growth factor *EG* as described in Hou et al. (2021). Using A and V to indicate cross-sectional average and variance, the table shows $A|a|$ and $A|t\,(a)|$, the averages across the 25 size-momentum portfolios of the absolute intercepts a and the absolute t-statistic for the intercepts; $Aa^2/V\bar{r}$, the average squared intercepts over the cross-section variance of $\bar{r}$ (the average returns on the 25 size-momentum portfolios); $A\lambda^2/V\bar{r}$, with $\lambda^2 \equiv a^2 - s^2(a)$, the average difference between each squared intercept and its squared standard error, $s^2(a)$, divided by the variance of $\bar{r}$; $A\bar{R}^2$, the average regression R^2; $As(a)$, the average standard error of the intercepts; $As(e)$, the average of residual standard deviations; $As^2(a)/A(a^2)$, the average of the estimates of the variances of the sampling errors of the estimated intercepts over $A(a^2)$; $Sh^2(a)$, the max. squared Sharpe ratio for the intercepts for the 25 size-momentum portfolios; the GRS statistic of Gibbons et al. (1989); and the p-value of GRS. The sample period is July 1972 to December 2016.

Model	$A\|a\|$	$A\|t\,(a)\|$	$Aa^2/V\bar{r}$	$A\lambda^2/V\bar{r}$	$A\bar{R}^2$	$As\,(a)$	$As\,(e)$	$As^2(a)/A(a^2)$	$Sh^2\,(a)$	GRS	$p(GRS)$
Panel A: Explaining excess returns of 5×5 size-momentum sorted portfolios											
MKTRF, ME, I/A, ROE	0.112	1.17	0.23	0.12	0.85	0.107	2.27	0.49	0.163	2.84	0.000
MKTRF, ME, I/A, ROE, EG	0.143	1.29	0.28	0.15	0.85	0.116	2.26	0.46	0.185	2.73	0.000
MKTRF, SMB*, HML*, RMW*, CMA**	0.402	1.78	2.37	1.87	0.33	0.233	4.85	0.21	0.215	3.61	0.000
MKTRF, SMB, HML, RMW, CMA	0.271	2.44	1.01	0.90	0.85	0.104	2.26	0.11	0.217	3.99	0.000
MKTRF, SMB, HML, RMW, CMA, UMD	0.122	1.57	0.21	0.15	0.92	0.077	1.65	0.26	0.186	3.32	0.000
MKTRF, SMB, HML, RMW, CMA, UMD^{RISK}	0.153	1.89	0.36	0.30	0.91	0.080	1.73	0.16	0.245	4.46	0.000
MKTRF, SMB, HML, RMW, CMA, UMD^{MISP}	0.203	1.95	0.58	0.49	0.89	0.096	1.97	0.16	0.224	3.61	0.000
Panel B: Explaining the mispricing-component in excess returns of 5×5 size-momentum sorted portfolios											
MKTRF, ME, I/A, ROE	0.151	4.76	0.61	0.60	0.58	0.028	0.59	0.02	0.805	14.03	0.000
MKTRF, ME, I/A, ROE, EG	0.076	2.32	0.15	0.13	0.61	0.029	0.56	0.11	0.143	2.12	0.001
MKTRF, SMB*, HML*, RMW*, CMA**	0.202	4.57	1.03	0.99	0.19	0.041	0.86	0.04	0.141	2.36	0.000
MKTRF, SMB, HML, RMW, CMA	0.207	6.79	1.12	1.11	0.61	0.027	0.59	0.01	0.178	3.26	0.000
MKTRF, SMB, HML, RMW, CMA, UMD	0.147	5.98	0.58	0.57	0.75	0.022	0.46	0.02	0.329	5.87	0.000
MKTRF, SMB, HML, RMW, CMA, UMD^{RISK}	0.184	6.54	0.89	0.88	0.67	0.025	0.53	0.01	0.095	1.72	0.017
MKTRF, SMB, HML, RMW, CMA, UMD^{MISP}	0.032	2.67	0.02	0.02	0.91	0.011	0.23	0.10	0.029	0.47	0.987
Panel C: Explaining the risk-component in excess returns of 5×5 size-momentum sorted portfolios											
MKTRF, ME, I/A, ROE	0.246	2.47	0.85	0.75	0.84	0.101	2.15	0.11	0.327	5.70	0.000
MKTRF, ME, I/A, ROE, EG	0.232	2.14	0.67	0.58	0.84	0.110	2.15	0.16	0.285	4.21	0.000
MKTRF, SMB*, HML*, RMW*, CMA**	0.487	2.29	3.02	2.61	0.35	0.213	4.45	0.14	0.253	4.25	0.000
MKTRF, SMB, HML, RMW, CMA	0.205	2.35	0.79	0.70	0.85	0.095	2.06	0.11	0.400	7.34	0.000
MKTRF, SMB, HML, RMW, CMA, UMD	0.260	3.21	0.85	0.79	0.90	0.079	1.71	0.07	0.411	7.33	0.000
MKTRF, SMB, HML, RMW, CMA, UMD^{RISK}	0.221	2.88	0.72	0.66	0.91	0.077	1.67	0.08	0.389	7.07	0.000
MKTRF, SMB, HML, RMW, CMA, UMD^{MISP}	0.235	2.34	0.73	0.64	0.87	0.097	1.97	0.12	0.316	5.11	0.000

Panel A of Table 4.5 confirms the well-known fact that FF5 leaves lots of momentum unexplained while HXZ outperforms. The estimates of squared intercepts over the cross-sectional variance of the average returns $Aa^2/V\bar{r}$ is 1.01, resp. 2.37 for the characteristic-efficient FF5 factors. These metrics are far above the values of this ratio for other models. Adding the UMD factor reduces $Aa^2/V\bar{r}$ to 0.21 but is achieved by explaining momentum by a coarser version of itself. The HXZ factors results in low average absolute intercepts of 0.112 that are statistically not distinguishable from zero. The most interesting part in Panel A is reported in the last two lines: How well do FF5 factors describe the 25 size-momentum portfolios when augmented with either the risk-, or the mispricing-component of UMD? First, both components improve the FF5 model by decreasing the absolute intercept from 0.271 to 0.153 (risk), resp. 0.203 (mispricing). Second, both of them result in insignificant absolute t-statistics of intercepts and the augmented models both have a higher value of 0.16 for $As^2(a)/A(a^2)$ than 0.11 for FF5, indicating that more of dispersion of intercept estimates is attributable to sampling error rather than to dispersion of the true intercepts.

In Panel B of Table 4.5, we test how well factor models describe only the mispricing-components among 25 size-momentum portfolios. As expected, all standard models perform worse compared to explaining compound portfolio returns. While the magnitude of average absolute intercepts remains approximately the same, they are estimated more precisely as indicated by remarkable high values of average absolute t-statistics, e.g., 6.79 for FF5 and 4.76 for HXZ. This is mostly because less than 11% of the dispersion of the intercepts is due to sampling error. FF5 augmented with the mispricing-component of UMD passes the GRS test (p-value: 0.987), so UMD^{MISP} successfully describes dispersion among the mispricing-component in 25 size-momentum portfolio returns. All metrics for FF5 and HXZ are quite similar but HXZ + EG outperforms all models that do not include a momentum factor. The average absolute intercept using the HXZ + EG model is only 7.6 bps and the results state that the models' superiority in explaining momentum is because it partially captures common mispricing in momentum returns. While our tests remain silent about the specific role of the ROE factor in explaining momentum (see Novy-Marx (2015)), we find that the expected investment growth factor EG extraordinary boosts the models' performance.

It is apparent from Panel C that both FF5 and HXZ fail to describe the risk-component in our 25 size-momentum portfolios, despite to their intent to capture exposure to risk. The average absolute intercept is within the range 0.205 to 0.487. The average standard error of intercepts is 0.095 for FF5 and 0.101 for HXZ and only up 16% of dispersion of the intercepts is due to sampling error. The explanatory power as measured by average R^2 in explaining the risk-component remains unchanged compared to Panel A.

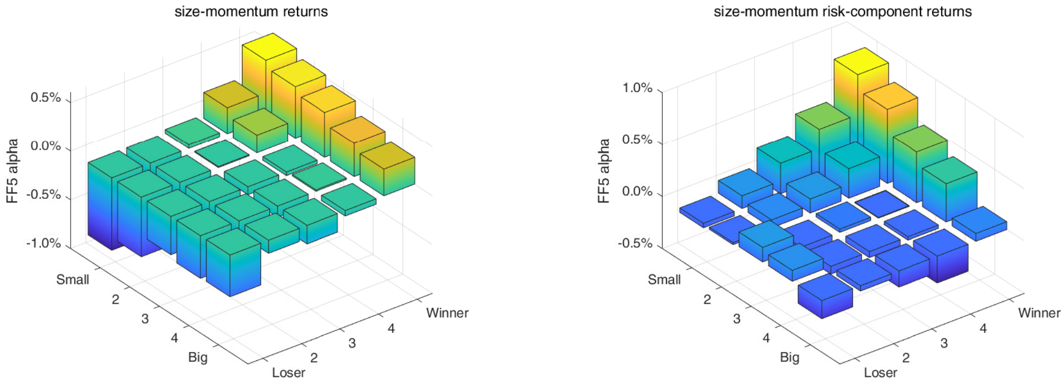

Fig. 4.2. This figure shows intercepts (alphas) from time-series regressions for 25 size-momentum portfolio returns on FF5 factors (left panel). The right panel shows according alphas for the risk-component in portfolio returns. The sample period is July 1972 to December 2016.

To our surprise, statistical metrics for the FF5 + UMD$^{\text{RISK}}$ augmented model are only slightly superior compared to FF5 + UMD. Fama and French (2016) demonstrate that FF5 leaves most of momentum returns among small firms unexplained. We confirm this finding and observe absolute t-statistics for alphas in extreme momentum quintiles among the lowest two size segments that are all larger than three. Similarly, FF5 fails to capture the return of all winner portfolios no matter the size quintile. The alpha for big winner is 0.27% with a t-statistic of 2.22 and increases to 0.60% (t-statistic: 4.84) for small winner. However, looking at the risk-components of our 25 size-momentum returns, a different picture emerges: FF5 now generates insignificant alphas for all loser portfolios. Even for the lowest size quintile, the alpha of the loser portfolio is an insignificant (t-statistic: -0.24) and economically negligible -0.04%. Another striking result to emerge from analyzing the risk-components is that average alphas for all winner portfolios apart from large winner tend to be a relatively large 0.50%. The alpha for large winner is an insignificant 0.07% but a remarkable 0.86% (t-statistic: 6.72) for small winner. Figure 4.2 summarizes these findings.

Overall, this section provides evidence that both FF5 and HXZ are incapable to capture even the risk-component in 25 size-momentum portfolio returns. Our findings reveal that high returns to small winners are the actual puzzle for FF5, whereas intercepts of risk-component returns for all loser portfolios are indeed small and insignificant.

4.4.4 Liquidity and volatility risk

After having isolated the risk-, and mispricing-component in momentum returns, we step into the ongoing debate about the relation between momentum, liquidity risk, and market volatility (see Avramov et al. (2016)) for one very specific reason: Focusing on these types of risk has the conceptual advantage that the economic source of risk being captured by related factors, namely liquidity and volatility, is well known.[18]

First, risk management techniques like volatility-scaling substantially improve momentum performance (see Hanauer and Windmüller (2023) and the literature therein). Second, momentum profits are remarkably larger in liquid market states, as documented in Lesmond et al. (2004) and Avramov et al. (2016). Sadka (2006) shows that a significant part of momentum returns is explained by a liquidity risk premium and Asness et al. (2013) extends these findings across international markets and different asset classes.

We contribute to this literature by focusing on their testable implications as follows. According to Huber (2022), the relation between between liquidity and mispricing can be either negative, or positive:[19] (i) In highly illiquid markets, strategies to exploit mispricing are difficult to implement. (ii) Highly liquid markets could potentially increase noise trading because of lower trading frictings and thus lead to higher mispricing. If the relation between momentum and liquidity is actually attributable to mispricing by one of these two channels, we should observe significant loadings in the mispricing-component of momentum returns. As for volatility-scaling, we would expect significant loadings only on the risk-component, because related strategies as described in Barroso and Santa-Clara (2015), Daniel and Moskowitz (2016), and Wang and Yan (2021) are the result of pure risk management.

We follow the empirical approach in Avramov et al. (2016) but consider both funding liquidity and market illiquidity (see Brunnermeier and Pedersen (2009)). Market illiquidity MKTILLIQ is defined as the value-weighted average of each NYSE/AMEX stock's monthly illiquidity. At the stock level, illiquidity is measured as in Amihud (2002), i.e., $\left[\sum_{d=1}^{n}|R_{i,d}|/\left(P_{i,d} \times N_{i,d}\right)\right]/n$, where n is the number of trading days in each month t, $|R_{i,d}|$ is the absolute value of the return of stock i on day d, $P_{i,d}$ is the daily closing price of stock i, and $N_{i,d}$ is the number of shares of stock i traded during day d. To proxy for

[18] The much broader question about what economic risks or state variables are reflected in the risk-component of momentum returns is addressed in Section 4.4.5.

[19] Similarly, Han et al. (2022) find that expected returns are related to trading volume positively among underpriced stocks but negatively among overpriced stocks, so they conclude that the volume-return relation is heterogeneous and depends on mispricing. As such, trading volume amplifies mispricing.

funding liquidity, we follow Asness et al. (2013) and employ the Treasury-Eurodollar (TED) spread, which is the average over the month of the daily 3-month interbank LIBOR interest rate (in U.S. dollar) minus the 3-month Treasury bill rate.[20] Aggregated market volatility MKTVOL is defined as the standard deviation of daily value-weighted U.S. equity market returns over a month. DOWN is a dummy variable that takes the value one if the past two-year cumulative value-weighted market return is negative and zero otherwise.

The average correlation between TED spread and MKTILLIQ is a modest 0.54 and in line with Wang and Xu (2015) and Avramov et al. (2016), UMD returns and one-month lagged MKTVOL are negatively correlated (-0.14). Table 4.6 depicts time-series regression estimates based on the following specification:

$$r_t = \alpha_0 + \beta_1 ILLIQ_{t-1} + \beta_2 DOWN_{t-1} + \beta_3 MKTVOL_{t-1} + c' F_t + \epsilon_t. \qquad (4.2)$$

The vector F stands for the FF5 factors. The dependent variable r_t is either the time-series of UMD returns, its according long- or short leg, or its risk- or mispricing-component. ILLIQ is either market illiquidity (MKTILLIQ, Panel A) or funding illiquidity (TED spread, Panel B).

One unanticipated result is that we do not detect a significant relation between illiquidity and UMD returns. To our surprise and in contrast to Avramov et al. (2016), the TED spread is even positively related with momentum returns. Further, the long leg of UMD is highly significantly related with both funding and market illiquidity. A one-standard deviation increase in MKTILLIQ (TED spread) corresponds with a return increase in the long leg of 0.23% (0.35%), which is also economically a non-negligible effect given that the average long leg UMD return is 1.34%. As with any findings in empirical research, this could be sample-specific within the meaning of Lo and MacKinlay (1990). First, Avramov et al. (2016) use a much longer sample period starting in 01/1928. Effective costs for NYSE/AMEX stocks exhibit considerable variation over time and highest costs are observed immediately after the 1929 crash and during the Depression (see Hasbrouck (2009)). Second, they analyze momentum decile 10-1 returns which are on average twice as large as UMD returns (see Table 4.3). Third, we additionally include the investment factor CMA and profitability factor RMW in our regressions to account for recent developments

[20] The LIBOR rate is taken from Refinitiv Datastream (USI60LDD) because the FRED series "TED Spread" is not available before 01/1986. The average absolute difference between the two time-series is 0.009 and their correlation exceeds 0.99.

Table 4.6: Momentum profits and market states.

This table shows results from regressing monthly UMD momentum returns (in %) on one-month lagged market illiquidity (value-weighted average of stock-level Amihud (2002) illiquidity, Panel A) or funding illiquidity (TED spread, Panel B), a dummy variable DOWN that takes the value of one if the cumulative 24-months market return is negative, and zero otherwise, the standard deviation of daily market returns (MKTVOL), and contemporaneous FF5 factors. To obtain the risk- and mispricing-component in UMD returns, we regress (using our baseline approach described in the main text) monthly value-weighted momentum portfolio returns on mispricing factors proposed in Stambaugh and Yuan (2017), Asness et al. (2019) and Daniel et al. (2020a), interaction terms indicating distinct market states (bear- and bull-market), and the market return as proposed in Daniel and Moskowitz (2016). The according fitted values correspond to the mispricing- and option-component, while the remainder (intercept, residuals, and market exposure) corresponds to the risk-component. Finally, we average each component across the three factor models. Newey and West (1987) corrected t-statistics are given in parenthesis and */**/*** indicate significance at the 10%/5%/1% level. The sample period is July 1972 to December 2016.

	UMD	UMD		UMD^{RISK}		UMD^{MISP}	
		Long	Short	Long	Short	Long	Short
Panel A: Time-series regressions on one-month lagged market illiquidity, market volatility, and controls							
Intercept	2.03***	1.02**	−1.01**	0.48***	−0.86***	0.10***	−0.48***
	(3.84)	(6.50)	(−2.52)	(3.97)	(−2.87)	(3.05)	(−3.96)
MKTILLIQ	1.85	2.11***	0.26	0.92	−0.46	0.20**	−0.68*
	(1.28)	(2.88)	(0.30)	(1.48)	(−0.67)	(2.14)	(−1.83)
DOWN	−1.26*	−0.63**	0.63	0.12	0.18	−0.04	0.36*
	(−1.83)	(−2.32)	(1.38)	(0.50)	(0.50)	(−0.99)	(1.92)
MKTVOL	−1.35**	−0.38**	0.98**	−0.26**	0.93***	−0.06	0.04
	(−2.40)	(−2.23)	(2.29)	(−2.03)	(2.83)	(−1.48)	(0.30)
MKTRF	−0.16*	1.00***	1.16***	1.03***	1.02***	0.00	0.11***
	(−1.92)	(35.18)	(20.02)	(43.03)	(24.31)	(0.63)	(6.44)
SMB	0.06	0.47***	0.41***	0.44***	0.38***	0.03***	0.04
	(0.52)	(11.35)	(4.74)	(11.80)	(6.01)	(4.13)	(1.34)
HML	−0.66***	−0.17***	0.49***	−0.06	0.26***	−0.09***	0.21***
	(−3.80)	(−3.11)	(3.82)	(−1.43)	(2.68)	(−9.16)	(6.32)
RMW	0.24	0.07	−0.17	0.08	0.18	−0.03**	−0.34***
	(1.03)	(0.85)	(−1.07)	(1.12)	(1.56)	(−2.12)	(−5.72)
CMA	0.55*	0.13	−0.42**	0.12	−0.15	0.00	−0.27***
	(1.95)	(1.33)	(−2.19)	(1.43)	(−1.07)	(0.03)	(−4.78)
Panel B: Time-series regressions on one-month lagged funding illiquidity, market volatility, and controls							
Intercept	1.96***	0.90***	−1.06***	0.44***	−0.83***	0.11***	−0.49***
	(3.63)	(5.72)	(−2.61)	(3.61)	(−2.74)	(3.16)	(−4.06)
TED spread	0.30*	0.40***	0.10	0.16*	−0.09	0.02	−0.06
	(1.69)	(4.74)	(0.92)	(1.95)	(−1.09)	(1.42)	(−1.38)
DOWN	−1.01*	−0.34	0.67*	0.25	0.12	−0.01	0.27
	(−1.66)	(−1.36)	(1.70)	(1.11)	(0.37)	(−0.31)	(1.51)
MKTVOL	−1.51**	−0.57***	0.94**	−0.34**	0.97***	−0.07*	0.08
	(−2.57)	(−3.17)	(2.12)	(−2.54)	(2.92)	(−1.70)	(0.59)
MKTRF	−0.15*	1.00***	1.16***	1.03***	1.02***	0.00	0.11***
	(−1.90)	(35.77)	(20.07)	(43.19)	(24.38)	(0.66)	(6.43)
SMB	0.06	0.48***	0.41***	0.44***	0.38***	0.03***	0.04
	(0.52)	(11.21)	(4.75)	(11.70)	(5.99)	(4.08)	(1.34)
HML	−0.66***	−0.17***	0.49***	−0.06	0.26***	−0.09***	0.21***
	(−3.79)	(−3.09)	(3.82)	(−1.42)	(2.68)	(−9.11)	(6.28)
RMW	0.24	0.07	−0.17	0.08	0.18	−0.03**	−0.34***
	(1.05)	(0.94)	(−1.05)	(1.16)	(1.56)	(−2.14)	(−5.73)
CMA	0.55*	0.13	−0.42**	0.12	−0.15	0.00	−0.27***
	(1.94)	(1.30)	(−2.20)	(1.41)	(−1.06)	(0.03)	(−4.75)

in asset pricing ahead in time of their study.[21] Another possible explanation is, that our option-component already accounts for some liquidity effects. This channel is suggested by Butt and Virk (2020) who document that variation in market liquidity is an important determinant of momentum crashes. When market liquidity increases in down market states then the returns of the loser portfolio increase more; an observation we account for with our option-component. Indeed, we detect that returns for the option-component are negatively correlated (-0.26) with one-month lagged market illiquidity. The according regression coefficient is a highly significant -1.30 (t-statistic: -2.50). We also apply (not tabulated) the liquidity measure proposed in Pastor and Stambaugh (2003) instead of MKTILLIQ for robustness of our results and still observe an insignificant relation between illiquidity and momentum returns.[22]

Our main result on illiquidity in this section is that we observe insignificant estimates for both the long-, and short leg of the risk-component. Exposure to illiquidity has the conceptual advantage that the source of the risk being captured (liquidity risk) is known. Our results imply that momentum returns are not a premium for bearing liquidity risk, because MKTILLIQ is unrelated with UMD^{RISK}. As for the mispricing-component UMD^{MISP}, we observe a significant positive coefficient of 0.20 (t-statistic: 2.14) for the long leg and a weakly significant negative coefficient of -0.68 (t-statistic: -1.83) for the short leg. Within the meaning of Huber (2022), we observe that the mispricing-component of momentum generates higher returns when the aggregated market is more illiquid. This supports the view that in highly illiquid markets, strategies to exploit mispricing are difficult to implement.

Our results quantitatively support the attempted, sketchy explanation in Asness et al. (2013) for the question why momentum loads positively on liquidity risk, while the value characteristic loads negatively on liquidity risk, and both factors generate a positive return premium.[23] Their simple and intuitive story is that momentum captures most popular trades, with investors buying assets whose prices have recently appreciated as fickle investors flocked to these assets. When a liquidity shock occurs, investor liquidations puts

[21] Analyzing the subperiod 04/2001 - 12/2011, we are able to replicate their results in Table 7 and observe a highly significant negative relation between one-month lagged MKTILLIQ and momentum (here: UMD) returns.

[22] The measure of Pastor and Stambaugh (2003) proxies for liquidity, hence, we take the negative of their measure so that it represents illiquidity.

[23] Our study fits their statement that *"further investigation into the opposite signed exposure of value and momentum to liquidity risk is an interesting research question, but beyond the scope of this paper"* (p. 962). More recently, Cooper et al. (2022) provide evidence that global macroeconomic risk exposure is likely to explain this puzzle.

more price pressure on these more crowded trades. In effect, Asness et al. (2013) suggest a mispricing-related explanation for the relation between illiquidity and momentum, and our results confirm that this relationship is prevalent only in the mispricing-component of momentum.

Next, focusing on market volatility, we document in line with the literature that momentum returns are significantly lower following periods of high market volatility. Looking at each momentum portfolio leg, winners negatively covary with market volatility which is further elaborated in a recent study (Misirli (2023)). On average, a one standard-deviation increase in MKTVOL corresponds to a decrease in UMD returns of 0.71% (Panel A), resp. 0.79% (Panel B). Because our estimates are very similar for market and funding illiquidity, we only describe results for Panel A (market illiquidity) in more detail. Momentum returns are 1.26 percentage points lower in times of bear market states. In line with Daniel and Moskowitz (2016), these phases of momentum crashes are captured by the (untabulated) option-component, which is the reason that DOWN is insignificant among the risk-, and mispricing-component.[24] The negative relation between momentum returns and the value factor HML is also in line with Avramov et al. (2016).

Volatility-scaling strategies as described in Barroso and Santa-Clara (2015), Daniel and Moskowitz (2016), and Wang and Yan (2021) are the result of pure risk management, so we expect significant MKTVOL loadings only on the risk-component. This is exactly what we observe: The long (short) leg of UMD^{RISK} is significantly exposed to aggregated market volatility risk with a coefficient of -0.26 (0.93). A one standard-deviation increase in MKTVOL lowers the return of the long-short risk-component by a highly significantly 0.62% (t-statistic: -2.76). As expected, UMD^{MISP} is unrelated with MKTVOL. Volatility-scaling momentum strategies generate very high returns, all among top six out of our 28 momentum strategies listed in Table 4.3. Our results confirm that their superiority is the outcome of effective risk management.

4.4.5 ICAPM interpretation for the risk-component of UMD

Analyzing the economic channels and associated economic risks driving momentum returns (at least partially) is far beyond the scope of this study. The previous section focuses especially on liquidity risk and market volatility and the reason is their conceptual

[24] We find that the estimated coefficient for DOWN using the option-component of UMD as dependent variable in Equation (4.2) is a highly significant -0.80 (t-statistic: -8.98). Both the long-, and short-leg are highly significantly negative, resp., positive, related with DOWN. This is as expected from our decomposition approach.

advantage of the source of risk, i.e., liquidity and volatility, being clearly identified. To shed more light on the economic interpretation of UMD's risk-component in this section, we interpret it as a pricing factor through the lens of the Intertemporal Capital Asset Pricing Model (ICAPM) in the style of Merton (1973), Petkova (2006), Welch and Goyal (2008), and Clarke (2022).

Following these studies, we begin with setting up a VAR model to capture the relation between macroeconomic state variables and the market return (both demeaned), as well as the predictable dynamics of each state variable:

$$\begin{bmatrix} MKT_t \\ DIV_t \\ TERM_t \\ DEF_t \\ RF_t \\ RISK_t \end{bmatrix} = A \begin{bmatrix} MKT_{t-1} \\ DIV_{t-1} \\ TERM_{t-1} \\ DEF_{t-1} \\ RF_{t-1} \\ RISK_{t-1} \end{bmatrix} + u_t. \tag{4.3}$$

The first term MKT_t is the excess market return and the remaining state variables are dividend-to-price ratio, the term spread, the default yield, the risk-free rate.[25] Additionally, the returns of the risk-component of UMD momentum ($RISK$) are included in the VAR system as potential state variables.

Following Petkova (2006), each series of innovations u_t is orthogonalized to the excess market return and scaled to match the variance of the market. These innovations act as unpredicted changes in respective state variables and are risk factors in addition to the excess return of the market portfolio. According to the ICAPM, exposures to these risk factors are important determinants of average portfolio returns, so we undertake cross-sectional asset pricing tests. Applying the two-pass procedure proposed in Fama and MacBeth (1973), the first-pass is a time series regression of portfolio excess returns $r^e_{i,t}$ on the market excess return and innovations in state variables:

[25]The dividend-to-price ratio (DIV) is the natural log of trailing sum of the 12-month dividends minus the natural log month-end value of the CRSP index. The term spread ($TERM$) is the yield on long-term U.S. Bonds minus the 3-Month Treasury Bill Secondary Market Rate (denoted as RF). The default spread (DEF) is the difference between the yield on BAA-rated corporate bonds and the long-term U.S. Bonds yield. These state variables are obtained from Amit Goyal's website (see Welch and Goyal (2008)).

$$r^e_{i,t} = \alpha_i + \beta_{i,MKT}MKT_t + \left(\beta_{i,\hat{u}^{DIV}}\right)\hat{u}^{DIV}_t + \left(\beta_{i,\hat{u}^{TERM}}\right)\hat{u}^{TERM}_t + \left(\beta_{i,\hat{u}^{DEF}}\right)\hat{u}^{DEF}_t + \left(\beta_{i,\hat{u}^{RF}}\right)\hat{u}^{RF}_t + \left(\beta_{i,\hat{u}^{RISK}}\right)\hat{u}^{RISK}_t + \varepsilon_{i,t}, \forall i. \tag{4.4}$$

We use the 25 size-momentum portfolios from Section 4.4.3 as test asset returns $r^e_{i,t}$. The second-pass involves relating the average excess returns of all assets to their exposures to the risk factors in the model:

$$\bar{r}^e_{i,t} = \gamma_0 + \gamma_{MKT}\hat{\beta}_{i,MKT} + \left(\gamma_{\hat{u}^{DIV}}\right)\hat{\beta}_{i,\hat{u}^{DIV}} + \left(\gamma_{\hat{u}^{TERM}}\right)\hat{\beta}_{i,\hat{u}^{TERM}} + \left(\gamma_{\hat{u}^{DEF}}\right)\hat{\beta}_{i,\hat{u}^{DEF}} + \left(\gamma_{\hat{u}^{RF}}\right)\hat{\beta}_{i,\hat{u}^{RF}} + \left(\gamma_{\hat{u}^{RISK}}\right)\hat{\beta}_{i,\hat{u}^{RISK}} + \epsilon, \forall t, \tag{4.5}$$

where γ represents the price of risk for innovations in the respective state variable, $\bar{r}^e_i$ is the average excess return on portfolio i, and ϵ is the residual. To account for the errors-in-variables problem when using the estimated betas from the time-series regression in Equation (4.4) as regressors in Equation (4.5), we follow the correction procedure in Shanken (1992). Allowing for a heteroskedasticity-robust inference, we also apply a GMM-procedure for computing standard errors. To account for possible serial correlation in GMM errors, we use the VARHAC method described in den Haan and Levin (2005) and Burnside (2011).[26]

Table 4.7 contains the results for Equation (4.5) that correspond to the second-pass of the Fama and MacBeth (1973) approach. To begin with, we use the 5×5 size-momentum portfolios as test assets in Panel A and B. Closely following the model presented in Petkova (2006), loadings on $\hat{u}^{RF}$ seem to be an important cross-sectional determinant of average UMD momentum returns.[27] The price of risk related with the term spread is weakly significant and the model leaves an economically large 1.36% p.m. left unexplained as indicated by the weakly significant intercept. Including innovations in the risk-component of UMD momentum returns (see Panel B) results in better performance metrics, e.g., the mean absolute pricing error reduces from 0.15 to 0.07 and the R^2 increases from 0.75 to 0.93. The significant loading on innovations in the risk-component of UMD momentum returns highlights that the risk-component is indeed an important component for explaining returns on portfolios formed on momentum. Turning to Panel C, we now use the according risk-component of each 5×5 size-momentum portfolio as test assets. The most important finding is the highly significant positive loading on $TERM$, while

[26] We thank Craig Burnside for making his code publicly available.

[27] These results are in line with a more recent study which finds evidence that macroeconomic shocks impact momentum strategy returns (see Sakemoto (2025)).

Table 4.7: Second-pass Fama and MacBeth (1973) regression estimates.

This table reports results for the second-pass Fama and MacBeth (1973) procedure, running cross-sectional regressions of mean excess returns on factor betas. The full-sample factor loadings, which are the independent variables in the regressions, are computed in one multiple time-series regression. The coefficients are expressed as percentage per month. We use the returns of 5 × 5 size-momentum portfolios as test assets in Panel A and B and their respective risk-component in Panel C. To obtain the risk-component, we use our baseline decomposition approach described in the main text (Section 4.3). The model includes the excess market return (MKT), and innovations in the dividend yield (DIV), term spread ($TERM$), default spread (DEF), the three-month Treasury Bill Secondary Market Rate (RF), and (only in Panel B and C) innovations in the returns of the risk-component of UMD momentum ($RISK$). For the factor risk premia, Shanken (1992) robust t-statistics are in parentheses, and GMM-VARHAC (den Haan and Levin (2005) and Burnside (2011)) robust t-statistics are in square brackets, and */**/*** indicate significance at the 10%/5%/1% level. The R^2 statistic from the second-pass regression is reported along with the mean absolute pricing error (MAE). The p-value correspond to F-statistics for the test that the pricing errors in the model are jointly zero. The sample period is July 1972 to December 2016.

Factor prices (λ)							R^2	p-value	MAE
Intercept	MKT	DIV	TERM	DEF	RF	RISK			
Panel A: 5 × 5 *size-momentum portfolios*									
1.36*	−0.56	−2.14	7.83*	−2.84	5.30***		0.75		0.15
(1.69)	(−0.72)	(−0.96)	(1.95)	(−0.86)	(2.88)			(0.77)	
[1.48]	[−0.60]	[−0.83]	[1.92]	[−0.78]	[2.44]			[0.75]	
Panel B: 5 × 5 *size-momentum portfolios*									
0.75	−0.07	−2.40	6.31*	3.13	5.74***	0.85**	0.93		0.07
(1.17)	(−0.10)	(−1.20)	(1.88)	(1.42)	(3.42)	(2.06)		(0.72)	
[1.41]	[−0.11]	[−1.33]	[1.80]	[1.48]	[3.44]	[2.15]		[0.66]	
Panel C: 5 × 5 *size-momentum portfolios - risk-component*									
−0.32	0.96	−2.65	8.52***	2.53	5.45***	0.42	0.91		0.08
(−0.31)	(0.94)	(−1.12)	(2.76)	(0.88)	(2.65)	(0.86)		(0.78)	
[−0.33]	[0.98]	[−1.23]	[2.80]	[0.84]	[2.58]	[0.85]		[0.74]	

innovations in the risk-component of UMD momentum returns become irrelevant in the model. The price of risk associated with term spread is 8.52, which means that if a portfolio has a unit beta with respect to $TERM$, this contributes 1.87% p.m. to the average excess return of that portfolio, given a one-standard deviation shock in $TERM$.[28]

Why is the risk-component of momentum related with term spread risk? In Section A.VI of the Appendix, we train a Gradient Boosted Trees (XGBoost) algorithm to identify relevant drivers of the risk-component of UMD among 126 U.S. macroeconomic variables and 43 anomaly portfolios unrelated with momentum strategies. We find that among the most important determinants are both the monthly rebalanced value strategy proposed by Asness and Frazzini (2013) and the traditional value strategy using annual book-to-market ratios. We find that high (traditional) value strategy returns correspond to high returns for the risk-component of UMD, while low value strategy returns give more rise for the importance of the rebalanced value strategy which are in turn negatively

[28]The volatility of innovations in $TERM$ is 4.55% p.m., so 8.52%/4.55% = 1.87%.

related with the risk-component. Overall, the literature vastly documents that value strategy returns are linked with term spread risk and variations over the business cycle (see Estrella and Hardouvelis (1991), Fama and French (1993), Fama and French (1995), Lettau and Ludvigson (2001) or Hahn and Lee (2006)), and our machine learning approach reveals that value returns serve as a channel that transmits term spread risk into the risk-component of momentum returns.[29]

In summary, the cross-section of raw 5×5 size-momentum returns is well captured by the risk-component of UMD returns and the risk-components of these 25 portfolios are strongly related with innovations in the term spread. These findings suggest, that the risk-component of UMD momentum returns proxies for a term spread surprise factor in returns, consistent with an ICAPM interpretation. In a more in-depth analysis presented in the Appendix, we show that value strategy returns are a key determinant for the risk-component of UMD returns and thus transmitting term spread risk into momentum.

4.5 Robustness tests and additional results

4.5.1 International evidence

Most of empirical research in asset pricing examines the U.S.-market (see Karolyi (2016)), so extending our findings to non-U.S. markets offers an important robustness test. Data on international mispricing factors MGMT, PERF, and QMJ, as well as international UMD and market returns returns are provided by Hanauer (2020) and Asness et al. (2019).[30] Because international data for PEAD and FIN is generally not available, we omit the mispricing model proposed in Daniel et al. (2020a) in our baseline decomposition approach.[31] We analyze the regions (i) global (developed) ex-U.S., (ii) Europe, and (iii) Japan, and all data is denominated in currency U.S. dollars. Because of limited data availability, the sample period starts in June 1992 for global ex-U.S., resp. June 1995 for Europa and Japan, and ends in June 2022 for all regions.

As in the United States, we find in Table 4.8 that all three components, i.e., risk, mispricing, and optionality, are important drivers of international momentum returns. In global (developed) ex-U.S. markets, the average UMD return is 0.59%, where 0.40% are attributable to the risk-component, 0.37% to the mispricing-component, and the

[29] In more recent studies, Liu and Moench (2016) and Moench and Stein (2025) show that the term spread is a robust out-of-sample predictor of recessions and market excess returns.

[30] `https://www.globalfactorpremia.org`.

[31] PEAD requires quarterly earnings announcement dates but reporting of quarterly data is not common practice in many non-U.S. countries before 2002.

Table 4.8: Risk-, mispricing-, and option-components of regional UMD momentum portfolios.

This table reports monthly average risk-, mispricing-, and option-components (in %) of UMD returns for the three regions global ex-U.S., Europe, and Japan. We regress (using our baseline approach described in the main text) monthly value-weighted momentum portfolio returns on mispricing factors proposed in Stambaugh and Yuan (2017) and Asness et al. (2019), interaction terms indicating distinct market states (bear- and bull-market), and the market return as proposed in Daniel and Moskowitz (2016). The according fitted values correspond to the mispricing- and option-component, while the remainder (intercept, residuals, and market exposure) corresponds to the risk-component. Finally, we average each component across the two factor models. Newey and West (1987) corrected t-statistics are given in parenthesis and */**/*** indicate significance at the 10%/5%/1% level.

			Monthly returns			
Region	Start	End	Avg.	Risk	Misp.	Opt.
Global ex-U.S.	06/1992	06/2022	0.59***	0.40***	0.37***	−0.19***
			(3.08)	(3.00)	(4.19)	(−4.76)
Europe	06/1995	06/2022	0.83***	0.54***	0.57***	−0.28***
			(3.34)	(3.01)	(4.32)	(−4.00)
Japan	06/1995	06/2022	0.07	−0.06	0.21*	−0.08***
			(0.27)	(−0.26)	(1.72)	(−4.57)

remaining -0.19% to the option-component. In Europe, the difference between risk-, and mispricing-component returns is also an economically unimportant 3 bps, highlighting their equal relevance. Consistent with Asness (2011) and the literature therein, the momentum strategy does not seem to work in Japan as indicated by an insignificant UMD return of 0.07%. Our results document only a weakly significant mispricing-component of 0.21%, and momentum crashes in the sense of Daniel and Moskowitz (2016) are also much lower in their economic magnitude compared to other regions. Overall, our results document a similar pattern in international UMD components as found in the U.S. market.

4.5.2 Transaction costs

Transaction costs always reduce the profitability of anomaly strategies while often being ignored. Lesmond et al. (2004) emphasize that momentum strategies require frequent trading in disproportionately high-cost securities. In effect, these trading costs prevent profitable strategy execution. However, Israel et al. (1999) analyze a momentum mutual fund launched in July 2009 by AQR Capital and conclude that according returns survive real-world trading frictions, taxes, costs, and expenses.

For robustness of our results, we decompose value-weighted long-short (decile 10-1) momentum returns net of transaction costs (measured by the effective bid-ask spread proposed by Hasbrouck (2009)) in this section. Data for the period 07/1972 - 12/2013 is provided by Novy-Marx's data library and discussed in Novy-Marx and Velikov (2016). To mitigate trading costs, the popular 10%/20% buy-and-hold rule is applied (see Arrow et al. (1951)), under which a trader will buy (sell short) stocks when they get into the top (bottom) 10% and hold them by restricting sales (short covers) only for stocks that leave the top (bottom) 20%. Related portfolio returns are constructed across market capitalization terciles determined by the 20th and 50th percentiles of NYSE stocks, so we separately analyze net momentum returns among microcaps, small, and big firms.

Table 4.9: Risk-, mispricing-, and option-components of momentum returns net of transaction costs.

This table reports monthly average risk-, mispricing-, and option-components (in %) of momentum returns net of transaction costs among three size segments microcap, small, and large, determined by the 20th and 50th percentiles of market capitalization of NYSE stocks. Trading costs are measured by the effective bid-ask spread proposed by Hasbrouck (2009). To mitigate trading costs, the popular 10%/20% buy-and-hold rule is applied (see Arrow et al. (1951)), under which a trader will buy (sell short) stocks when they get into the top (bottom) 10%, and hold them by restricting sales (short covers) only for stocks that leave the top (bottom) 20%. We decompose momentum returns using our baseline approach described in the main text (Section 4.3). Newey and West (1987) corrected t-statistics are given in parenthesis and */**/*** indicate significance at the 10%/5%/1% level. The sample period is July 1972 to December 2013.

	Gross	Net returns			
	Avg.	Avg.	Risk	Misp.	Opt.
Microcap	1.94***	1.14***	0.52**	0.90***	−0.28***
	(5.37)	(3.21)	(2.05)	(6.23)	(−3.60)
Small	1.31***	0.82**	0.23	0.79***	−0.19***
	(4.06)	(2.57)	(0.95)	(6.33)	(−3.60)
Large	0.81***	0.46	−0.08	0.60***	−0.06***
	(2.60)	(1.47)	(−0.30)	(5.93)	(−3.60)

The 10%/20% buy-and-hold momentum strategy generates a highly significant return difference among large winners and large losers of 0.81% as shown in Table 4.9. Considering microcaps, the strategy generates a huge return difference of 1.94%. Accounting for transaction costs which are on average 0.80%, 0.49%, and 0.35% for microcaps, resp. small, and large firms, renders the strategy for large firms insignificant. Despite the drop in returns between microcaps and large firms, the mispricing-component remains quite strong with average returns between 0.60% (large) and 0.90% (microcaps). In consequence, the plunge in momentum returns is driven by the risk-component which decreases from a significantly 0.52% for microcaps to an insignificant -0.08% for large firms. Similarly, microcaps are more prone for momentum crashes which is reflected in a huge option-component of -0.28%.

Why is the risk-component negligible for large and small firms but not microcaps? Applying the regression setting stated in Equation (4.2) reveals that the risk-component among large firms negatively covaries with market volatility as indicated by our highly significantly estimated coefficient $\hat{\beta}_3 = -2.14$ (t-statistic: -2.87). While the same holds for small firms ($\hat{\beta}_3 = -2.00$ (t-statistic: -2.25)), the risk-component is unrelated with market volatility for microcaps ($\hat{\beta}_3 = -0.97$ (t-statistic: -1.00)). Volatility risk, however, can be avoided as successfully demonstrated by volatility-scaling momentum strategies and the according exposure is unpriced (see Section 4.4.4).[32]

4.5.3 Daily return analysis

For further robustness of our main results, we apply our (baseline) decomposition approach using daily returns. This offers the possibility to reexamine the findings in Birru (2018) for decomposed momentum returns. The study reveals that anomalies for which the speculative leg, i.e., the portfolio holding stocks that are hard to arbitrage or that are highly subjective to value, is the short leg, experience highest returns on Mondays. The author links this pattern to investor mood which tends to decrease on Monday and increases from Thursday to Friday (see e.g., Rossi and Rossi (1977) and Stone et al. (2012)). These findings provide another important testable implication for the validity of our empirical methods: Because investor mood and mispricing-related returns positively covary and mood generally peaks on Fridays, we expect returns for the mispricing-component to be higher at the end of the week compared to Mondays. Again, we assume that mispricing generally takes the form of overvaluation instead of undervaluation, so we hypothesize according return patterns to arise mostly from the short leg of the mispricing-component, inversely driving overall momentum returns related with common mispricing.

The average daily UMD return between July 1972 and December 2016 is a highly significant 2.87 bps (t-statistic: 3.40) and contains (i) a risk-component of 1.50 bps (t-statistic: 2.25); (ii) a mispricing-component of 1.63 bps (t-statistic: 6.06); and (iii) an option-component of -0.26 bps (t-statistic: -5.55). The relative importance of the components is similar with our findings using monthly returns. Figure 4.3 shows that daily UMD returns are a highly significant 4.54 bps on Mondays, resp. 4.84 bps on Tuesdays, and tend to be weaker at the end of the week.

[32] Using the risk-components of 25 size-momentum portfolios discussed in Section 4.4.3 further confirm these findings. The coefficient estimate for one-month lagged MKTVOL is -1.34 (t-statistic: -2.43) for momentum among the highest size quintile portfolios but only -0.93 (t-statistic: -1.70) among the lowest size quintile portfolios.

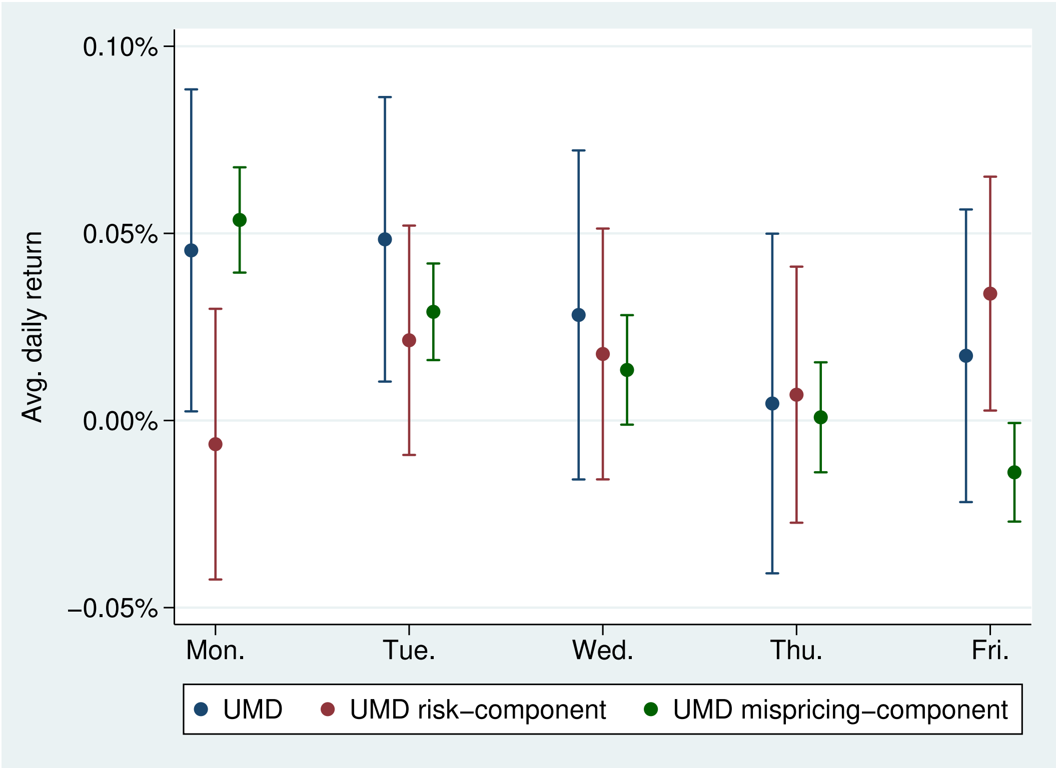

Fig. 4.3. This figure shows average daily returns for momentum (UMD) and its according risk-, and mispricing-component (using the baseline decomposition approach described in Section 4.3). The whiskers correspond to a 99% confidence interval according to Newey and West (1987) robust standard errors. The sample period is July 1972 to December 2016.

We find supporting evidence for the validity of our decomposition approach within the context of Birru (2018) because the mispricing-component is strongly related with the days of the week: The mispricing-component of UMD generates high returns of 5.36 bps on Mondays which plunge to a significantly negative -1.38 bps until Fridays. This pattern is entirely rooted in the short leg of the momentum strategy where the mispricing-component contributes -6.29 bps on Mondays (t-statistic: -11.24) and 1.74 bps on Fridays (t-statistic: 3.48), inversely driving returns of the overall momentum portfolio strategy. As for the long leg of momentum, the mispricing-component accounts for less than one basispoint in absolute values (-0.93 bps on Mondays and 0.36 bps on Fridays) among all trading days which seems to be economically negligible.

4.5.4 Dynamic momentum (mis)pricing

How accurately is our decomposition approach presented in Section 4.3 in separating momentum returns into a risk-, mispricing-, and option-component? This section provides a novel robustness test based on the insights of van Binsbergen et al. (2023) who estimate the dynamics of price wedges, i.e., dislocations of the (absolute) price level, for well-known anomaly portfolios. They define price wedges PW_t as the deviation of the stock's market price P_t from its informationally efficient (fundamental) value $\tilde{P}_t$, which is defined as the present value of future dividends $\{D_s\}_t^\infty$ and capital gains under a benchmark SDF $\frac{m_{t+s}}{m_t}$:

$$PW_t = -\log\left(\frac{\tilde{P}_t}{P_t}\right) = -\log\left(\mathbb{E}\left[\sum_{s=1}^{J} \frac{m_{t+s}}{m_t}\frac{D_{t+s}}{P_t} + \frac{m_{t+J}}{m_t}\frac{P_{t+J}}{P_t}\right]\right). \tag{4.6}$$

They set $J = 180$ assuming that any price wedges have converged to zero 15 years after the initial portfolio formation and estimate price wedges at the portfolio level by taking the average of price wedges across the portfolio cohorts. Their candidate SDF is an exponentially affine representation of the single-factor CAPM such that the price wedge of the market portfolio equals zero. van Binsbergen et al. (2023) conclude that momentum is a build-up anomaly, i.e., abnormal expected momentum returns further exacerbate price dislocations.[33] Further, at the time of portfolio formation, the momentum portfolio is already mispriced, and subsequent momentum returns exacerbate these existing price dislocations. All price dislocations stem from the first decile portfolio (loser) having a highly significant estimated price wedge of -21.6%, indicating that loser stocks are already underpriced at the time of the portfolio formation. Winner (decile 10) are, however, accurately priced with an insignificant price wedge of only 3.5%.

It is important to notice that this method does not impose perfect foresight by investors but rather is consistent with the premise that fundamental stock valuations $\tilde{P}$ are based on rational expectations. For that reason, we calculate relative changes in $\tilde{P}$ which we label as *rational* returns r^{rat} in the meaning that they capture changes in fundamental stock prices which are based on rational expectations.[34] We can't simply take the authors original time-series of $\tilde{P}_t$ (at the portfolio level) because of monthly portfolio formations, so $\tilde{P}_{t+1}$ would refer to portfolio cohorts in $t+1$. These cohorts in $t+1$ vastly differ from cohorts in t because of the high turnover of momentum strategies. For that reason, we re-implement their approach to accurately calculate r_t^{rat} for the loser and winner decile portfolio based on according initial portfolio cohorts.[35]

[33] We use their value-weighted decile 10-1 portfolios sorted on cumulative past 12-month performance skipping the most recent month as momentum strategy in this section.

[34] We thank the authors for making their code and data publicly available.

[35] To be more precise, we calculate the expectation in Equation (4.6) starting in month $s = 2$ until $J = 181$ for portfolios formed in $t = 0$ and use the SDF to shift resulting present values into $t+1$. This results in an estimate of $\tilde{P}_{t+1}$ for the initial portfolio cohorts.

Taken together, the procedure outlined in van Binsbergen et al. (2023) enables us to evaluate the robustness and validity of our decomposition approach by analyzing the following testable implications:

a) Momentum's long leg (decile 10) is accurately priced at the time of portfolio formation, so according rational returns $r_t^{rat,L}$ are expected to be free of a mispricing-component. Deviations from actual momentum long leg returns $r_t^L - r_t^{rat,L}$ should be indistinguishable from zero.

b) Momentum's short leg (decile 1) is substantially undervalued at the time of portfolio formation and subsequent momentum returns exacerbate this price dislocation (see van Binsbergen et al. (2023)). According rational returns $r_t^{rat,S}$ are thus expected to contain a significant mispricing-component which should be positive to not only reflect the initial undervaluation but also the subsequent push to even lower prices. Deviations from actual momentum short leg returns $r_t^S - r_t^{rat,S}$ are expected to be negative to account for this mispricing.

c) The construction of rational momentum returns r_t^{rat} is based on a CAPM-based SDF, so according alphas are expected to equal zero.

Table 4.10 summarizes our decomposition results rational momentum returns over the period July 1972 to Nov. 2002.[36]

Panel A reexamines our analysis of decomposing momentum decile portfolio returns into a risk-, mispricing-, and option-component for the shorter sample period ending in Nov. 2002. The according long-short strategy generates a highly significant return of 1.58% p.m. with a CAPM alpha of 1.63%. More interesting, Panel B shows decomposition results for rational momentum returns r_t^{rat}. The average rational return of the long leg $r_t^{rat,L}$ is on average 1.29% and we estimate its according mispricing-component to be an economically negligible 0.10%. The risk-component accounts for an economically large 1.05% while being statistically insignificant. Regressing this risk-component of $r_t^{rat,L}$ on FF5, HXZ, or simultaneously on all mispricing factors (DHS, SY, AS) factors reveals that it entirely captures market risk and does not covary with other factors. Because the long leg is accurately priced at the portfolio formation date, we are not able to observe significant differences in neither the risk-, nor the mispricing-component between actual and rational returns (see Panel C).

[36]Data ends in Nov. 2002 because Equation (4.6) requires buy-and-hold dividends and capital gains up to 15 years after portfolio formation.

Table 4.10: Risk-, mispricing-, and option-components of rational momentum returns.

This table reports monthly average risk-, mispricing-, and option-components (in %) of value-weighted decile momentum portfolio excess returns (Panel A) and according rational excess returns (Panel B). Rational returns are relative changes in rational prices, which are calculated as the present value of future dividends and capital gains over the next 15 years (see van Binsbergen et al. (2023)). We use an exponentially affine CAPM SDF that sets the price wedge, i.e., deviations of actual market prices from their rational price, for the aggregate market equal to zero. To obtain decomposed momentum returns, we use our baseline approach described in the main text (Section 4.3). The sample period is July 1972 to November 2002. */**/*** indicate significance at the 10%/5%/1% level according to Newey and West (1987) corrected standard errors.

	Monthly returns				t-statistic				CAPM alpha	
	Avg.	Risk	Misp.	Opt.	Avg.	Risk	Misp.	Opt.	Avg.	t-stat.
Panel A: Value-weighted momentum decile excess returns										
Long	1.60***	1.40***	0.10**	0.10***	4.67	4.46	2.03	2.66	1.06***	6.41
Short	0.02	0.97***	−0.91***	−0.04***	0.05	2.97	−8.12	−2.66	−0.56**	−2.43
Long-Short	1.58***	0.44	1.00***	0.14***	4.88	1.46	8.75	2.66	1.63***	5.07
Panel B: Rational value-weighted momentum decile excess returns										
Long	1.29*	1.05	0.10*	0.14***	1.91	1.52	1.95	2.66	0.15	0.36
Short	1.84*	0.59	0.68***	0.57***	1.92	0.59	8.57	2.66	0.65	1.09
Long-Short	−0.55	0.45	−0.58***	−0.43***	−0.68	0.54	−7.33	−2.66	−0.50	−0.66
Panel C: Decile minus rational returns										
Long	0.31	0.36	−0.01	−0.04***	0.56	0.66	−0.08	−2.66	0.92	1.81
Short	−1.82**	0.37	−1.58***	−0.61***	−2.23	0.42	−8.43	−2.66	−1.22	−1.77
Long-Short	2.13**	−0.02	1.58***	0.57***	2.27	−0.02	8.55	2.66	2.13**	2.37

On the other hand, the undervalued short leg yields 1.84% and our decomposition approach is able to detect its underpricing by extracting a highly significant 0.68% (t-statistic: 8.57) mispricing-component. In line with a mispricing explanation, rational returns of past loser are significantly higher compared to actual short leg returns (the difference $r_t^S - r_t^{rat,S}$ is a significant -1.82%, see Panel C) and $r_t^{rat,S}$ positively covaries with the $PERF$ factor while negatively loading on HML.

The analysis presented in this section vastly differs from previous robustness tests because the return series under investigation is specifically constructed to reflect rational expectations based on the CAPM. Overall, our results are in line with van Binsbergen et al. (2023) and demonstrate that our decomposition approach successfully extracts mispricing even in these rationally derived momentum returns.

4.5.5 Simulation evidence regarding the decomposition approach

In this section, we assess statistical properties and the power of our decomposition approach by means of bootstrap simulations. Due to the construction of our decomposition approach stated in Equation (4.1), the market return collapses into a pure risk-component. We analyze a total of 534 monthly returns and start to randomly replace the market returns by mispricing returns. Since the one extreme of 0% replacements is the case of decomposing the market return and the other extreme of 100% replacements equals decomposing pure mispricing returns, we are able to evaluate how well our approach detects mispricing. We repeat each simulation 250 times and display in Fig. 4.4 the average return together with the according mean of its mispricing-component and the mean of the mispricing-components' t-statistic. The shaded area covers the 99% confidence interval based on bootstrapped standard errors using the method of Politis and Romano (1994) with 100 simulation runs where the optimal block-length is estimated using the method of Politis and White (2004) and Patton et al. (2009).

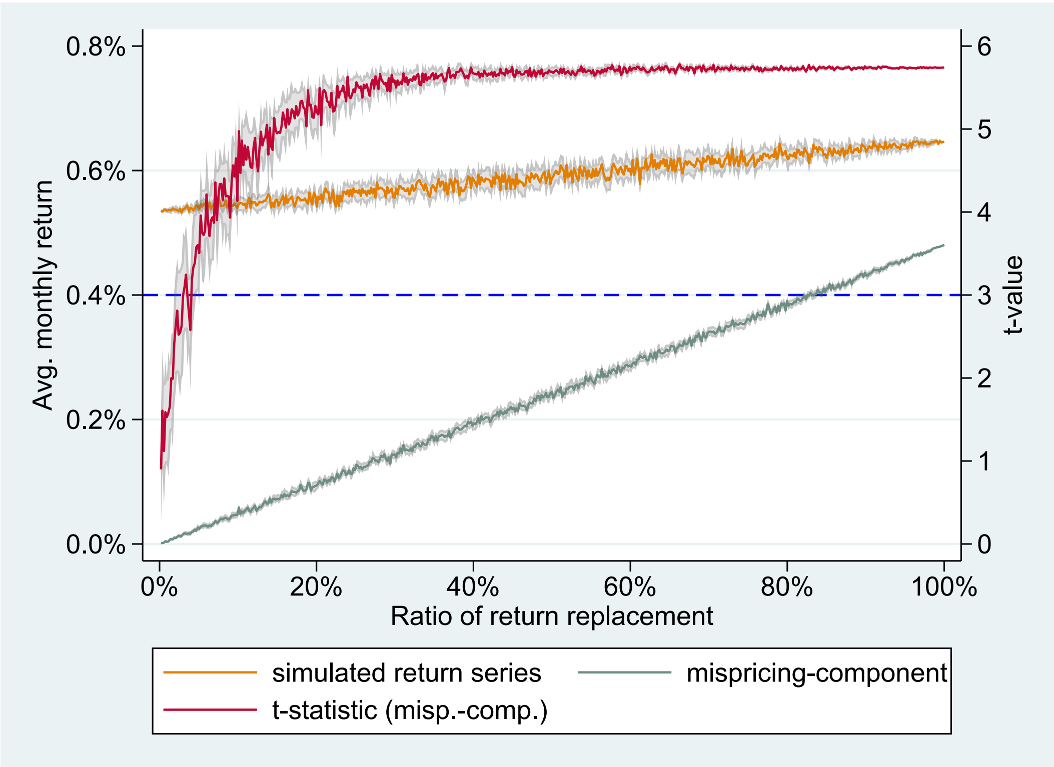

Fig. 4.4. This figure shows average simulated returns where monthly market excess returns are gradually replaced with mispricing returns, calculated as the first pricincipal component of mispricing factor returns (DHS, SY, and AS). We decompose the simulated returns into its according risk-, and mispricing-component (using the baseline decomposition approach described in Section 4.3). The gray shaded area corresponds to a 99% confidence interval based on bootstrapped standard errors using the method of Politis and Romano (1994) with 100 simulation runs where the optimal block-length is estimated using the method of Politis and White (2004) and Patton et al. (2009). We repeat each simulation 250 times. The sample period is July 1972 to December 2016.

The mispricing returns we use as gradually replacements for market returns stem from the first principal component of mispricing factor returns (DHS, SY, and AS). As suggested in Campbell et al. (1997), we scale each factor by the sum of the loadings, so that the weights sum to one and are as follows: 0.33 for *FIN*, 0.23 for *PERF*, 0.20 for *MGMT*, 0.18 for *QMJ*, and 0.05 for *PEAD*. This results in a time-series of mispricing returns generating on average 0.65% p.m. and having an annualized Sharpe ratio of 0.91. For that reason, the average return for simulated returns varies between 0.54% (market exc. return, resp. 0% replacement) and 0.65% (mispricing return, resp. 100% replacement of market returns) and the according mispricing-component detected varies between 0.00% and 0.48%. This demonstrates that our decomposition approach is quite conservative in extracting the mispricing-component, as it only attributes 74.39% of the first principal component of mispricing factor returns to be mispricing. Similarly, our method detects on average only a conservative 69.52% mispricing-component among mispricing factor returns themselves (between 41.10% for *PEAD* and 85.56% for *PERF*). The reason for this is that we apply Equation (4.1) for each of our three mispricing models and then average the risk-, mispricing-, and option-components across the three factor models. The blue dashed-line refers to the threshold value of three for the Newey and West (1987) robust t-statistic which is generally attained after replacing only 21 out of 534 (i.e., 3.93%) market returns with corresponding mispricing returns. At this ratio, the average mispricing-component is 0.02% which accounts for only 3.03% of the average simulated return (0.54%). Overall, these results show that our proposed method has strong statistical power.

4.6 Conclusion

In this paper, we decompose the returns of 28 U.S.-equity momentum strategies into a risk-, mispricing-, and option-component using a novel decomposition approach and provide a direct test for various explanations. Two distinct pillars for the momentum effect emerge from the vast body of literature: Risk and behavioral explanations. Our paper contributes to bridge this gap and mitigate the academic dispute as our findings imply that momentum returns are essentially driven by both, mispricing and risk. We are the first to quantify their respective relevance and document that standard *UMD* momentum returns of 0.66% p.m. contain (i) an insignificant 0.29% risk-component, (ii) 0.50% mispricing-component, and (iii) -0.13% option-component (accounting for momentum crashes in the sense of Daniel and Moskowitz (2016)). Looking at all 28 different variants of momentum strategies, we find that the risk-component accounts on average for 56.66% of momentum returns, followed by a 29.63% mispricing-component, and a remaining 13.71% option-component.

Investor sentiment induced mispricing varies over time and conclusions drawn from empirical studies generally depend on the specific sample period (Lo and MacKinlay (1990)). Take e.g., contrary results for the relation between the default spread between U.S. corporate bonds and momentum returns. Chordia and Shivakumar (2002) report a strong negative relation while Griffin et al. (2003) fail to confirm this evidence. We detect that the relation is almost entirely driven by the risk-component of the short momentum portfolio leg. Separating risk- and mispricing in momentum seems to be crucial for our understanding of this phenomenon, especially since decades of intensive research did not result in a commonly accepted explanation. The relation between momentum and market illiquidity as documented in Avramov et al. (2016) is fully subsumed by the mispricing-component of momentum. This finding supports the view in Huber (2022) that strategies to exploit mispricing are difficult to implement in illiquid markets. The negative relation between momentum and aggregated market volatility is fully reflected in its risk-component. Momentum returns are, at least partially, an actual risk premium for market volatility risk, which explains the superior performance of volatility-scaling momentum strategies proposed in Barroso and Santa-Clara (2015), Daniel and Moskowitz (2016), and Wang and Yan (2021). Tests of standard models in asset pricing can be compromised if loadings on systematic mispricing are correlated with risk (see Birru et al. (2023)). We reexamine if the Fama/French five-factor model (Fama and French (2015)) and the q-factor model (Hou et al. (2015)) are able to describe the risk-component of 25 size-momentum portfolios. The answer is no. Average absolute intercepts among the risk-components are larger in magnitude and estimated more precisely compared to composed momentum returns. This finding even holds if we augment the aforementioned models with the time-series return of the risk-component. Our findings suggest that innovations in the term spread carry an important price of risk and have important pricing implications for the cross-section of portfolios sorted on momentum, and value strategy returns seem to be a key determinant for the risk-component of momentum. However, much work is left to identify how exactly value strategy returns transmit term spread risk into the risk-component of momentum. In effect, after controlling for the mispricing-component, the "risk" in momentum seems to be an even greater puzzle than previously thought.

Appendix for
"The Relevance of Risk, Mispricing, and Optionality in Momentum Returns"

Abstract

- Section A.I: Robustness tests for the baseline decomposition approach.
- Section A.II: Principal component analysis of the baseline approach.
- Section A.III: IPCA factor decomposition.
- Section A.IV: Time-series evolution of UMD momentum components.
- Section A.V: Spanning regressions.
- Section A.VI: A machine-learning approach (XGBoost) to identify macroeconomic and financial drivers of the risk-component of momentum returns.

A.I Robustness tests for the baseline decomposition approach

Our baseline decomposition approach for decomposing momentum returns is presented in Section 4.3. In this section, we use different variants of this approach presented in Equation (4.1) to evaluate the robustness of our results with respect to changes in the baseline model. To begin with, subsequent sections present variations of the baseline approach. Then, we provide an overview in form of presenting coefficient estimates for our decomposition models using momentum decile 10-1 (L-S) returns in Section A.I.

Orthogonalized mispricing factors

The mispricing factors in Equation (4.1) are constructed using portfolio sorts and are not required to be orthogonal on market returns. We observe an average correlation of -0.38 between mispricing factors and the market excess return, so any mispricing factor could unintentionally capture a portion of market risk. We account for that possibility and orthogonalize each of the mispricing factors to the market factor as a robustness check:

$$X_{j,t} = \alpha_{i,j} + \beta_{i,j} r^e_{m,t} + \epsilon_{i,j,t}. \tag{4.7}$$

Specifically, we regress each factor return $X_{j,t}$ on the market excess return and use the sum of estimated intercept and residuals $\hat{\alpha}_{i,j} + \epsilon_{i,j,t}$ as orthogonalized mispricing factors $X_{j,t}^{\perp MKT}$ in Equation (4.1).

Quantile regression decomposition

Another caveat may be that our baseline regression model estimates the conditional mean of return components. This may not be favorable if the impact of regressors may differ over the distributional range or if the response variable is biased because of extreme observations and outliers. The mean of L-S momentum returns is on average 1.21% while the median is 1.71%, highlighting an economically huge difference of 0.50% as shown in Table 4.1. This is attributable to the negative skewness (-1.44) of their distribution which is also highly leptokurtic having an excess kurtosis of 7.37. While the fat-tailed distribution of stock returns is widely known for a long time in the literature (see Mandelbrot (1963) or Fama (1965) for one of the earliest studies), the higher occurrence of extreme returns (esp. negative returns) and their influence in econometric models is not commonly addressed in empirical research.

To address this issue, we apply quantile regressions proposed by Koenker and Bassett

(1978) and Koenker (2005). Given our baseline model from Equation (4.1), the estimated intercepts $\hat{\alpha}_{i,j}$ and coefficients $\hat{\beta}_{i,j}$ are now dependent on the parameter τ and estimate the (conditional) τ-quantile of momentum return components. Assuming $Q_\tau(\epsilon_{i,j,t}(\tau)) = 0$, i.e., errors $\epsilon_{i,j,t}(\tau)$ are uncorrelated and allowed to be highly heterogeneous, the skewed and fat-tailed distribution of returns is accounted for. The τ-specific estimates of $\hat{\alpha}_{i,j}(\tau)$ and $\hat{\beta}_{i,j}(\tau)$ are obtained by minimizing the following objective function:

$$\sum_{t=1}^{T} \rho_\tau \left(r^e_{i,t} - \left(\alpha_{i,j} + \alpha^B_{i,j} \cdot I_{B,t-1}\right) - \left(\beta^0_{i,j} + I_{B,t-1}\left(\beta^B_{i,j} + I_{U,t} \cdot \beta^{B,U}_{i,j}\right)\right) r^e_{m,t} - \beta_{i,j} X_{j,t} \right)$$
$$\text{s.t.}$$
$$\rho_\tau(\omega) = \begin{cases} \tau\omega & \text{, if } \omega \geq 0, \\ (1-\tau)|\omega| & \text{, else.} \end{cases} \tag{4.8}$$

Equation (4.8) minimizes residuals that are weighted by an asymmetric loss function. This provides an unbiased and consistent estimator for the conditional τ-th quantile of return components corresponding to mispricing model j. We set $\tau = 0.50$ to estimate the conditional median in comparison to the (implicitly) conditional mean estimation in our baseline model.

Ridge regression decomposition

The average correlation between mispricing factors and the market excess return is -0.38 so our model in Equation (4.1) could be ill-posed because of correlated regressors. We resort to regularization and estimate our model using ridge regression shrinking estimated coefficients towards zero. If we use the notion for the error term from our baseline decomposition model

$$\epsilon_{i,j,t} \equiv r^e_{i,t} - \left(\alpha_{i,j} + \alpha^B_{i,j} \cdot I_{B,t-1}\right) - \left(\beta^0_{i,j} + I_{B,t-1}\left(\beta^B_{i,j} + I_{U,t} \cdot \beta^{B,U}_{i,j}\right)\right) r^e_{m,t} - \beta_{i,j} X^S_{j,t},$$

the estimator for the penalized ridge regression model is given by

$$\min_{\alpha_{i,j},\beta_{i,j} \in \mathbb{R}} \left\{ \frac{1}{T} \sum_{t=1}^{T} \epsilon^2_{i,j,t} \right\} \quad \text{s.t.} \quad \mathbb{1}^T \beta^2_{i,j} \leq q,$$

according to factor model j, or in Lagrangian form:

$$\min_{\alpha_{i,j},\beta_{i,j}\in\mathbb{R}} \left\{ \frac{1}{T}\sum_{t=1}^{T} \epsilon_{i,j,t}^2 + \lambda \min \|\beta_{i,j}\|_2^2 \right\}. \tag{4.9}$$

The parameters $q \geq 0$ and L_2 regularization $\lambda \geq 0$ refer to as shrinkage parameters. The penalty term λ results in biased coefficient estimates but the rationale of the procedure is that the reduction in the mean squared error of the model often outweighs the size of the bias. We select λ using 10-fold cross validation for the mean squared error criterion as generally recommended in Arlot and Celisse (2010) and Bergmeir and Benítez (2012).

Two-stage approach

Daniel and Moskowitz (2016) argument that momentum strategies in bear markets (i.e., $\prod_{t-24}^{t-1}(1 + r_{m,t}) \leq 1$) behave like written call options on the market and thus lose much when the market rises (i.e., $r_{m,t}^e > 0$). These exceptional high returns for loser stocks are, however, only rarely observed. On average, the excess return for the loser portfolio is -0.19% but conditional on the 45 out of 534 observed months where past loser embody the option premium, they generate an astonishing 9.75% compared to -1.11% otherwise. This non-linear and only temporarily observed option premium could bias our estimates for conditional means in Equation (4.1).

To address this issue, we apply a two-stage approach as an additional robustness test. First, we use the original formula proposed in Daniel and Moskowitz (2016),

$$r_{i,t}^e = (\alpha_0 + \alpha_B \cdot I_{B,t-1}) + (\beta_0 + I_{B,t-1}(\beta_B + I_{U,t} \cdot \beta_{B,U}))\, r_{m,t}^e + \epsilon_{i,t}, \tag{4.10}$$

where $r_t^{OPT} \equiv \hat{\beta}_{B,U} I_{B,t-1} I_{U,t} r_{m,t}^e$ denotes the (directly estimated) option component in momentum portfolio returns $r_{i,t}^e$. In consequence, r_t^{OPT} does not depend on a certain mispricing model anymore. Second, we apply the original method in Birru et al. (2023) to dissect option-adjusted momentum returns into a risk- and mispricing-component:

$$r_{i,t}^e - r_t^{OPT} = \alpha_{i,j} + \beta_{i,j} X_{j,t} + \epsilon_{i,t}, \tag{4.11}$$

where $\hat{\beta}_{i,j} X_{j,t}$ is the mispricing component and the remainder attributes to the risk component, both according to factor model j.

Decomposition using risk factors

Decomposing momentum returns based on mispricing factors is central to our analysis. Our framework relies on the assumption that the return component unexplained by mispricing factors primarily reflect non-behavioral forces. The residual component (denoted as the risk component) could still include further effects not captured by the mispricing factors, but making use of three distinct models mitigates concerns of data mining and any portion of systematic mispricing in the residual component labeled as "risk" should at least be substantially weaker. As a robustness test, we estimate Equation (4.1) redefining $X_{j,t}$ to be a vector of risk factors corresponding to factor model FF5, HXZ, and BS. We drop the momentum factor *UMD* in BS to avoid the circular reasoning of explaining momentum by momentum, but add either the liquidity factor *LIQ* or the betting-against beta factor *BAB*.[37] Nevertheless, adding any (pure) risk factor in this robustness test for the baseline decomposition approach only further weakens our results and thus provides a more cautious framework.

Panel regression decomposition

Using decile momentum portfolios has the advantage to provide an in-depth analysis of more granular portfolios but neglects that nearby portfolio returns strongly covary as indicted by the very high average correlation of 0.91. Standard bivariate portfolio sorts commonly reflect that covariation by applying 30% and 70% percentile breakpoints and assigning all stocks in according extreme portfolios to an anomaly under study. We reflect that by using a panel regression approach: The ten decile portfolios are divided into three panels according to portfolios 1-to-3, 4-to-7, and 8-to-10. We apply Equation (4.1) with portfolio fixed effects and standard errors clustered by portfolio and month. To provide further evidence for the robustness of our results, we use market-orthogonalized mispricing-factors while dropping *MKTRF* from the model to eliminate any direct measurement of risk.

Bayesian change-point decomposition

Finally, we use a Bayesian framework for our baseline decomposition model applying the Markov Chain Monte Carlo (MCMC) method of Chib (1998) to generate a sample from the posterior distribution (see Geman and Geman (1984), Gelfand and Smith (1990),

[37]The choice of replacing *UMD* by one of these factors is somehow arbitrary, however, both are widely used in the literature and are more clearly linked to capture risk exposure, namely illiquidity and market risk, compared with other possible factors.

Kass and Raftery (1995), Chib (1995)). The centerpiece of this approach is that we allow the model from Equation (4.1) to contain structural changes in the distribution of momentum returns.[38] Applying a regime switching model matches the tendency of momentum returns to change their behavior abruptly as seen during momentum crashes.

Let $Y_t := (r^e_1, r^e_2, ..., r^e_T)$ denote the monthly time-series of $T = 534$ momentum portfolio excess returns (for ease of notation, we drop the portfolio indicator i), and let $p\,(y_t|Y_{t-1}, \xi_t)$ be the conditional distribution of y_t with time-specific parameters $\xi_t = (\xi_{1t}, ..., \xi_{pt})$. ξ_t changes at latent time points $\tau_1, ...\tau_m$ so that $\xi_t = \theta_j$ for all $t \in [\tau_{j-1}, \tau_j)$ and $\Theta := (\theta_1, ..., \theta_{m+1})$. Suppose $s_t \in \{1, 2, ..., m+1\}$ denotes the latent state variable at time t, the random sequence $s_1, s_2, ..., s_m$ is assumed to be a Markov chain with transition probabilities $p\,(s_{t+1} = j+1|s_t = j) = 1 - p\,(s_{t+1} = j|s_t = j)$. In detail, $s_t = j$ denotes that the structural parameter at time t is attributed to regime j. The modeling of the distinct values θ_j, for $j = 1, ..., m$, is given by a prior distribution F. Then, the likelihood of a model having M states is given by

$$
\begin{aligned}
p\,(Y_T|\Theta) &= p\,(y_1|\xi_1, s_t = 1) \prod_{t=2}^{T} \sum_{m=1}^{M} p\,(y_t|\xi_m, s_t = m)\, p\,(s_t = m|Y_{t-1}, \Theta)\,, \\
\xi_t &= \begin{cases} \theta_1 & \text{, if } t < \tau_1, \\ \theta_2 & \text{, if } \tau_1 \leq t < \tau_2, \\ \vdots & \\ \theta_{m+1} & \text{, if } \tau_m \leq t \leq T, \end{cases} \\
\theta_j|F &\sim F,\; j = 1, ..., m+1.
\end{aligned}
\tag{4.12}
$$

For our model, we need to estimate the joint distribution of the parameter set Θ, the transition probability matrix $P = \{p_{i,j}\}_{1 \leq i,j \leq s}$ with $p_{ij} = p\,(s_t = j|s_{t-1} = i)$, and the posterior probabilities for the regime sequence $s_1, s_2, ..., s_m$. P is completely determined by its diagonal elements, so one has only to estimate $p_{s,s}$. Without loss of generality and for ease of notation, our baseline model in Equation (4.1) takes the form $r^e_{(T\times 1)} = X_{(T\times k)}\beta_{(k\times 1)} + \epsilon_{(T\times 1)}$, with $\epsilon \sim \mathcal{N}(0, \sigma^2 I_{T\times T})$, having k regressors (incl. intercept) and I as the identity matrix. Adopting the Bayesian paradigm, we use the following standard prior distributions:

[38] Commonly, structural breaks are accounted for by analyzing sub-periods, discarding data from the sample outside the specified time interval. In addition, this also assumes that the exact timing of structural breaks is observable.

$$\beta_s \sim \mathcal{N}_k(b_0, B_0^{-1}), \quad \sigma_s^2 \sim \mathcal{IG}(\frac{c_0}{2}, \frac{d_0}{2}), \quad p_{s,s} \sim \mathcal{B}(x_0, z_0), \\ s = 1, ..., m+1, \tag{4.13}$$

where $\mathcal{N}$, $\mathcal{IG}$, and $\mathcal{B}$ is the normal, inverse gamma, and beta distribution, respectively, and B_0 the prior precision of b_0. As recommended in Chib (1998), x_0 and z_0 are set to match the expected duration, i.e., the sample period ($T = 534$ months) divided by the number of a-priori fixed states $M \in \{0, ..., 5\}$. Similar to Gelman (2006), we use an uninformative prior for the variance of residuals by setting ($c_0 = d_0 = 0.001$). To facilitate model comparison, we need at least a weakly informative prior for the distribution of β_s because marginal likelihoods are not identified under diffuse priors. Theoretically, a coefficient $\gamma_{i,t}$ of regressing stock excess returns $r_{i,t}^e$ on a given factor f_t is bounded by $\left[-\frac{\sigma(r_{i,t}^e)}{\sigma(f_t)}, \frac{\sigma(r_{i,t}^e)}{\sigma(f_t)}\right]$. At the portfolio level subsuming N stocks with market capitalization as weight $w_{i,t}$, we observe $\gamma_{p,t} = \sum_{i=1}^{N} w_{i,t}\gamma_{i,t}$, which is tilted towards one in case of well diversified portfolios with large N.[39] Empirically, if we regress the 55 long-short anomaly portfolio returns used in Kozak et al. (2018) and Haddad et al. (2020) on each of the mispricing factors in DHS, SY, and AS, we observe coefficient estimates within the interval $[-1.26; 1.86]$ with 99% confidence. For that reason, the commonly used weakly informative prior $b_0 = 0$ (i.e., the mispricing factor is irrelevant) with a precision $B_0 = 0.1$ is not only plausible, but rather conservative.

We detect the number of structural changes m by comparing candidate models with varying, a-priori fixed numbers of structural breaks using (log) Bayes factors denoted as B. Two models $\mathcal{M}_a$ and $\mathcal{M}_b$ with a-priori a and b structural breaks are evaluated by the difference of log marginal likelihoods, i.e., $\ln(B_{a,b}) := m(Y_T|\mathcal{M}_a) - m(Y_T|\mathcal{M}_b)$. For $\ln(B_{a,b}) > 0$, we conclude that the data supports $\mathcal{M}_a$ more than $\mathcal{M}_b$ and vice versa (Jeffreys (1961)). Marginal likelihoods are calculated using the method of Chib (1995) that relies on the output of the Gibbs sampling algorithm. We run the algorithm for posterior sampling for a total number of $N = 2,000$ iterations, of which the first $N/2$ are discarded as burn-in.[40]

Revisiting our baseline decomposition approach using the Bayesian framework presented in this section starts with detecting the appropriate number of structural change-points. We a-priori assume up to five regime changes in decile portfolio 10-1 momentum returns

[39] This collapses to the fact that the value-weighted average of the market beta considering all securities must equal one in case of f being the market factor.

[40] Results are similar for $N = 10,000$ iterations, of which the first 3,000 are discarded as burn-in.

Table A.I: Bayes factors for the Bayesian change-point model.

Bayes factors (on the natural log scale) for the Bayesian change-point analysis of momentum decile portfolio 10-1 returns from July 1972 to December 2016. Each model with a-priori $m \in \{1, 2, 3, 4, 5\}$ change-points is compared to the identical model but assuming zero change-points as benchmark. The first column indicates which mispricing model is used for the Bayesian regression estimation. The second column shows the date of the estimated change-point according to $m = 1$. Marginal likelihoods are calculated using the method of Chib (1995) and we run the algorithm for posterior sampling for a total number of $N = 2,000$ iterations, of which the first $N/2$ are discarded as burn-in.

		Number of change-points				
Mispricing model	Change-point τ_1	$m = 1$	$m = 2$	$m = 3$	$m = 4$	$m = 5$
Stambaugh and Yuan (2017)	08/1998	11.14	−1.40	−14.48	−27.56	−39.19
Daniel et al. (2020a)	07/1998	12.07	−6.00	−17.25	−27.40	−22.61
Asness et al. (2019)	06/1998	10.13	5.13	−0.42	−7.49	−9.40

and measure the strength of evidence for that model compared to a model without any change-points. Table A.I reports related (log) Bayes factors $B_{m,0}$, with $m \in \{1, 2, 3, 4, 5\}$.

Our Bayesian analysis strongly supports a model with one change-point occurring in mid of 1998 in momentum long-short portfolio returns, no matter which specific mispricing model is used. To be specific, the posterior probability that a model having one change-point is correct given that one of the models under study is correct, is 99.77% on average, so we choose $m = 1$ to decompose momentum returns in our Bayesian setting. As in our baseline approach, the mispricing component of portfolio returns is the sum of the product(s) of the estimated loading(s) and factor(s), but loadings now depend on the state s_t prevailing in t.

Comparison of the baseline model with different decomposition approaches

This section starts with presenting coefficient estimates for our decomposition models using momentum decile 10-1 (L-S) returns from Kenneth French's website in Table A.II.[41]

[41] We do not report coefficient estimates for the baseline model with market-orthogonalized mispricing factors because the estimate of β^0 is only slightly affected. The panel regression approach is only meaningful for groups of momentum portfolios and therefore not reported here. We also set aside explaining momentum by risk factors which is already vastly documented in the literature (see e.g., Fama and French (1996) or, more recently, Kelly et al. (2021b).

Table A.II: Coefficient estimates for long-short momentum returns.

This table reports coefficient estimates of monthly value-weighted long-short (decile portfolio 10-1) momentum returns for our decomposition models using mispricing-factors proposed in Stambaugh and Yuan (2017) (SY), Asness et al. (2019) (AS), and Daniel et al. (2020a) (DHS). The first column refers to the estimation method and the second column indicates the according mispricing model. The last column reports R^2 adjusted for the degrees of freedom. Robust t-statistics are presented in parenthesis and */**/*** indicate significance at the 10%/5%/1% level. We use Newey and West (1987) corrected standard errors with a lag of six months for the baseline model and the Two-stage approach. The quantile model's t-statistics are based on the Markov chain marginal bootstrap of He and Hu (2002) using 600 iterations. We do not report t-statistics for ridge regression estimates because it still remains an unresolved issue, resp. there is no consensus in the literature, how to calculate them (see Kyung et al. (2010)). Time-series robust standard errors for the Bayesian analysis are reported in square brackets (in quantity 10^{-2}) and are based on an estimate of the spectral density at zero using 1,000 draws from the posterior distribution of each coefficient. The sample period is July 1972 to December 2016.

Estimation method	Model	Regime	α	α^B	PEAD	FIN	MGMT	PERF	QMJ	β^0	β^B	$\beta^{B,U}$	σ^2	Adj. R^2
Baseline	DHS		0.26	−0.77	1.64***	0.19				0.07	−0.54**	−0.26		0.27
			(0.76)	(−0.57)	(7.67)	(1.02)				(0.75)	(−2.48)	(−0.49)		
	SY		0.37	−1.20			0.30*	1.30***		0.22***	−0.25	−0.21		0.52
			(1.54)	(−0.97)			(1.94)	(14.82)		(3.43)	(−0.85)	(−0.36)		
	AS		1.00***	−0.35					0.91***	0.18*	−0.54**	−0.40		0.16
			(2.72)	(−0.26)					(2.99)	(1.90)	(−2.11)	(−0.63)		
Quantile	DHS		0.55*	0.04	1.57***	0.10				−0.03	−0.38	−0.40		
			(1.90)	(0.03)	(9.34)	(0.79)				(−0.40)	(−1.05)	(−0.54)		
	SY		0.56**	−0.95			0.42***	1.23***		0.19**	−0.15	−0.02		
			(2.03)	(−0.77)			(2.69)	(11.29)		(2.30)	(−0.44)	(−0.04)		
	AS		1.36***	0.00					0.74***	0.06	−0.53	−0.16		
			(5.67)	(0.00)					(4.48)	(0.81)	(−1.31)	(−0.18)		
Ridge	DHS		0.39	−0.52	1.49	0.16				0.03	−0.44	−0.36		
	SY		0.49	−1.02			0.24	1.20		0.16	−0.23	−0.26		
	AS		1.13	−0.34					0.72	0.08	−0.42	−0.43		
Two-stage	DHS		0.06		1.75***	0.24								0.22
			(0.15)		(6.65)	(1.47)								
	SY		0.44*				0.15	1.28***						0.49
			(1.65)				(1.07)	(13.15)						
	AS		0.94**						0.98***					0.10
			(2.50)						(3.06)					
Bayes	DHS	07/1972 - 07/1998	0.29	−1.84	1.25	0.38				0.18	−0.36	0.30	24.45	
			[1.14]	[4.72]	[0.62]	[0.38]				[0.26]	[0.95]	[1.39]	[7.23]	
	DHS	08/1998 - 12/2016	0.44	1.95	1.68	0.06				−0.09	−0.28	−1.51	50.49	
			[1.67]	[4.96]	[0.71]	[0.36]				[0.45]	[0.99]	[1.66]	[14.06]	
	SY	07/1972 - 08/1998	0.48	−3.29			0.68	1.31		0.22	−0.73	0.96	17.55	
			[0.95]	[3.62]			[0.44]	[0.30]		[0.25]	[0.87]	[1.24]	[5.15]	
	SY	09/1998 - 12/2016	−0.13	1.41			0.08	1.36		0.36	−0.00	−1.33	30.20	
			[1.22]	[4.15]			[0.40]	[0.27]		[0.36]	[0.75]	[1.31]	[8.49]	
	AS	07/1972 - 06/1998	1.43	−1.42					0.62	0.14	−0.56	0.43	29.31	
			[1.13]	[5.12]					[0.66]	[0.27]	[1.10]	[1.49]	[8.15]	
	AS	07/1998 - 12/2016	0.67	1.97					0.88	0.18	−0.33	−1.62	57.36	
			[1.60]	[4.83]					[0.70]	[0.50]	[1.04]	[1.78]	[15.46]	

Newey and West (1987) t-statistics using a lag of six are reported for the baseline model and the Two-stage approach. T-statistics for the quantile model are based on robust standard errors using the Markov chain marginal bootstrap of He and Hu (2002). We do not report t-statistics for ridge regression estimates because it still remains an unresolved issue, resp. there is no consensus in the literature, how to calculate them (Kyung et al. (2010)). To retrieve time-series robust standard errors for coefficients in the Bayesian framework, we draw 1,000 samples from the posterior distribution for each coefficient c and denote the according vector of samples $s_{c,i}$. Using the naive MCMC error, i.e., the volatility of drawn samples, disregards the potential autocorrelation of the MCMC samples. For that reason, we apply an autoregressive model over $s_{c,i}$, selecting the order K by the Akaike information criterion. If the fitted model is

$$s_{c,i} - \bar{s}_{c,i} = \sum_{k=1}^{K} \rho_k \left(s_{c,i} - \bar{s}_{c,i}\right) + \epsilon_{c,i}, \quad \epsilon_{c,i} \sim \mathcal{N}(0, \sigma_\epsilon^2), \tag{4.14}$$

then the time-series robust standard error se_c according to coefficient c is obtained by

$$se_c = \sqrt{\frac{\sigma_\epsilon^2}{\left(1 - \sum_{k=1}^{K} \rho_k\right)^2}}. \tag{4.15}$$

Our baseline approach reveals that momentum L-S returns have a highly significant exposure to more transient ($PEAD$) and short-term ($PERF$) mispricing, while persistent (FIN) and long-term mispricing ($MGMT$) are negligible. Momentum returns also significantly covary with the returns of the QMJ factor while market risk exposure is remarkable low with an average $\hat{\beta}^0$ of only 0.16. In line with Daniel and Moskowitz (2016), exposure to market risk ($\beta^0 + \beta^B$) becomes negative when the past 24-month cumulative market return is negative. Otherwise, we are not able to detect a significant additional decline in market betas when the market return suddenly upswings during bear markets, although signs for according full-sample coefficient estimates are all negative. Considering that our Bayesian change-point analysis suggests one structural breakpoint around mid-1998 and reapplying the baseline model only for the second regime (08/1998 - 12/2016) results in an average $\hat{\beta}^{B,U}$ of -1.72, significant among all three mispricing models. Dropping all mispricing factors and applying the original model of Daniel and Moskowitz (2016) stated in Equation (4.10) results in a very similar estimate of -1.98. This indicates that the option-like behavior of momentum L-S returns is almost orthogonal to common mispricing.

To evaluate the robustness of our baseline decomposition approach, average risk-, mispricing-, and option-components of value–weighted decile momentum portfolio excess returns using all decomposition variants presented in this section are reported in Table A.III.

All our alternative decomposition approaches shown in Panel B to Panel H broadly support our previous findings presented in Table 4.2. The magnitude of the mispricing-component is on average 180% of the risk-component among all estimation methods and the former is estimated much more precisely as indicated by t-statistics above 3.80. In all models, the mispricing-component exceeds the risk-component.

Table A.III: Risk-, mispricing-, and option-components of momentum sorted decile portfolios.

This table reports monthly average risk-, mispricing-, and option-components (in %) of value-weighted decile momentum portfolio excess returns (first row). In Panel A, we regress momentum returns on mispricing factors proposed in Stambaugh and Yuan (2017), Asness et al. (2019), and Daniel et al. (2020a), interaction terms indicating distinct market states (bear- and bull-market), and the market return, as proposed in Daniel and Moskowitz (2016) and Birru et al. (2023). The according fitted values correspond to the mispricing-, and option-component, while the remainder (intercept, residuals, and market exposure) corresponds to the risk-component. Finally, we average each component across the three factor models (DHS, SY, AS). Panel B uses market-orthogonalized mispricing factors. Panel C to Panel H show decomposition results based on estimation methods described in Section A.I of the Appendix to provide further robustness of our baseline decomposition approach. The sample period is July 1972 to December 2016. */**/*** indicate significance at the 10%/5%/1% level according to Newey and West (1987) corrected standard errors.

	Decile Portfolio										L-S	
	1 Short	2	3	4	5	6	7	8	9	10 Long	Mean	t-stat.
Avg. return	−0.19	0.34	0.45*	0.59***	0.48**	0.54***	0.58***	0.72***	0.74***	1.01***	1.21***	3.62
Panel A: Baseline decomposition												
Risk	0.47	0.64**	0.52**	0.61***	0.46**	0.56**	0.40**	0.48**	0.58**	0.95***	0.49**	1.97
Mispricing	−0.77***	−0.37***	−0.18***	−0.06***	−0.00	0.07***	0.18***	0.19***	0.20***	0.07**	0.85***	6.32
Option	0.11***	0.07***	0.11***	0.04***	0.03***	−0.09***	0.01***	0.05***	−0.05***	−0.01***	−0.13***	−3.57
Panel B: Baseline decomposition using market-orthogonalized mispricing factors												
$\text{Risk}^{\perp MKT}$	0.63	0.70**	0.53**	0.60***	0.44**	0.54**	0.35	0.43**	0.54**	0.97***	0.35	1.35
$\text{Misp.}^{\perp MKT}$	−0.93***	−0.43***	−0.20***	−0.05**	0.01	0.09***	0.23***	0.24***	0.25***	0.06	0.99***	8.68
Option	0.11***	0.07***	0.11***	0.04***	0.03***	−0.09***	0.01***	0.05***	−0.05***	−0.01***	−0.13***	−3.57
Panel C: Quantile regression decomposition												
Risk	0.41	0.69***	0.58**	0.66***	0.46**	0.53**	0.36	0.49**	0.58**	1.00***	0.60**	2.41
Mispricing	−0.71***	−0.36***	−0.18***	−0.07***	0.01	0.08***	0.18***	0.18***	0.21***	0.08	0.78***	6.48
Option	0.11***	0.01***	0.04***	−0.00***	0.01***	−0.07***	0.03***	0.04***	−0.06***	−0.07***	−0.17***	−3.57
Panel D: Ridge regression decomposition												
Risk	0.44	0.64**	0.56**	0.64***	0.51***	0.57***	0.47**	0.58***	0.64***	1.03***	0.60**	2.38
Mispricing	−0.75***	−0.38***	−0.22***	−0.10***	−0.06***	0.00	0.10***	0.11***	0.13***	0.01	0.76***	6.28
Option	0.12***	0.09***	0.10***	0.05***	0.03***	−0.04***	0.01***	0.02***	−0.03***	−0.03***	−0.15***	−3.57
Panel E: Two-stage decomposition approach												
Risk	1.14***	1.27***	1.10***	1.19***	1.02***	1.18***	0.98***	1.06***	1.22***	1.62***	0.48	1.94
Mispricing	−1.45***	−1.00***	−0.76***	−0.64***	−0.56***	−0.55***	−0.41***	−0.39***	−0.43***	−0.59***	0.86***	6.30
Option	0.11***	0.07***	0.11***	0.04***	0.03***	−0.09***	0.01***	0.04***	−0.05***	−0.02***	−0.13***	−3.57
Panel F: Risk factor decomposition approach												
Risk	0.28	0.51**	0.62***	0.66***	0.65***	0.72***	0.72***	0.77***	0.83***	0.66***	0.38**	2.34
Mispricing	−0.60***	−0.24*	−0.29***	−0.10	−0.19***	−0.09	−0.14**	−0.11	−0.05	0.33***	0.93***	3.80
Option	0.13***	0.08***	0.12***	0.04***	0.03***	−0.09***	−0.00***	0.05***	−0.04***	0.02***	−0.11***	−3.57
Panel G: Panel regression decomposition												
Risk	0.31	0.85***	0.95***	0.51**	0.40*	0.46**	0.50**	0.48**	0.50**	0.78***	0.47**	2.09
Mispricing	−0.57***	−0.57***	−0.57***	0.11***	0.11***	0.11***	0.11***	0.30***	0.30***	0.30***	0.88***	5.50
Option	0.07***	0.07***	0.07***	−0.03***	−0.03***	−0.03***	−0.03***	−0.07***	−0.07***	−0.07***	−0.14***	−3.57
Panel H: Bayesian change-point decomposition												
Risk	0.26	0.62**	0.57**	0.65***	0.51**	0.64***	0.43**	0.51**	0.59**	0.96***	0.71***	2.93
Mispricing	−0.73***	−0.41***	−0.27***	−0.12***	−0.06**	0.01	0.15***	0.18***	0.21***	0.10**	0.83***	6.35
Option	0.28**	0.13***	0.14**	0.06**	0.03	−0.11***	0.00	0.03	−0.07***	−0.05	−0.33**	−2.01

A.II Principal component analysis of the baseline approach

As described in Section 4.3, we decompose momentum returns in our baseline approach with respect to a given mispricing factor model. Our procedure starts with regressing momentum returns on the mispricing factor(s) proposed in Stambaugh and Yuan (2017), Asness et al. (2019), and Daniel et al. (2020a). We conduct separate regressions for each of these three models. Fitted values correspond to the mispricing- and option-component, while the remainder (intercept, residuals, and market exposure) corresponds to the risk-component (see Equation (4.1). In a last step, we average each component across the three factor models. In this section, we aim to mitigate concerns with this last step.

Using principal component analysis, we extract the first principal component of mispricing factor returns. As suggested in Campbell et al. (1997), we scale each factor by the sum of the loadings, so that the weights sum to one and are as follows: 0.33 for *FIN*, 0.23 for *PERF*, 0.20 for *MGMT*, 0.18 for *QMJ*, and 0.05 for *PEAD*. The first two components explain 85.89% of the variance of the factor returns, 53.68% and 32.20%, respectively. This implies, that mispricing factors tend to be highly correlated.

To evaluate the robustness of our decomposition approach, instead of averaging component returns across the three mispricing factor models, we extract principal components of all mispricing factors and use the first component as the single, latent factor for our decomposition approach:

Step 1: Extract the first principal component r_t^{PCA} of all mispricing factors *MGMT*, *PERF*, *FIN*, *PEAD*, and *QMJ*.

Step 2: Set $X_{j,t} = r_t^{PCA}$, i.e., the first principal component acts as a single mispricing factor in Equation (4.1).

Step 3: Extract mispricing-, risk-, and option-component as shown in Equation (4.16).

Using the first principal component of mispricing factor returns allows us to directly extract the mispricing-, risk-, and option-component and avoids to average components across each separate model estimate:

$$r_{i,t}^e = \underbrace{\beta_{i,j} r_t^{PCA}}_{\text{mispricing}} + \underbrace{\beta_{i,j}^{B,U} I_{B,t-1} I_{U,t} r_{m,t}^e}_{\text{option}} + \underbrace{\left(\alpha_{i,j} + \alpha_{i,j}^B \cdot I_{B,t-1}\right) + \left(\beta_{i,j}^0 + I_{B,t-1}\beta_{i,j}^B\right) r_{m,t}^e + \epsilon_{i,j,t}}_{\text{risk}}, \tag{4.16}$$

We decompose all 28 momentum strategies shown in Table 4.3 using this PCA based approach and compare our results with the baseline approach. Table A.IV shows related differences in estimated component returns.

We observe that differences in average returns of the risk- and mispricing-component are statistically indistinguishable from zero for 20 out of 28 momentum strategies. This implies that taking the average of a component return across the three mispricing factor models is valid and our main results are not affected by this decision. Looking at UMD momentum returns, both approaches extract a highly significant mispricing-component of 0.50% p.m. and the risk-component only differs by 1 basispoint. We do, however, find significant differences among PC factor momentum as proposed in Ehsani and Linnainmaa (2022), but these differences are economically unimportant as they are less than 2 bps in magnitude. Notable differences are observed in momentum strategies which also tilt towards value investments, e.g., the combined value/momentum strategy in Asness et al. (2013) or junk stock momentum in Avramov et al. (2007). In these cases, averaging momentum components across our three mispricing models slightly underestimates the risk-component and overestimates the magnitude of the mispricing-component compared with the PCA-based approach. In the worst case, junk momentum (Avramov et al. (2007)) yields on average 11.88% p.a. and our baseline approach attributes 9.72% p.a. to risk, whereas the PCA-based approach concludes that the risk-component contributes a total of 11.88% to overall junk momentum returns (the mispricing- and option-component yield 3.54%, resp. -3.55%, and cancel out each other).

Table A.IV: Comparing risk- and mispricing-components of equity momentum strategies.

This table reports monthly average risk- and mispricing-components (in %) of 28 equity momentum strategies presented in Table 4.3 using our baseline decomposition approach described in Section 4.3 in comparison with a PCA-based approach. In the baseline approach, we (separately) regress monthly value-weighted momentum returns on mispricing factors proposed in Stambaugh and Yuan (2017), Asness et al. (2019), and Daniel et al. (2020a), interaction terms indicating distinct market states (bear- and bull-market), and the market return as proposed in Daniel and Moskowitz (2016). Following Birru et al. (2023), the according fitted values correspond to the mispricing- and option-component, while the remainder (intercept, residuals, and market exposure) corresponds to the risk-component. Finally, we average each component across the three factor models. Using the PCA-based decomposition approach for robustness of our results, we first extract principal components of all mispricing factors and use the first component as a single (mispricing) factor in Equation (4.1). Applying the PCA-based approach thus avoids to average components across each of the three separate (mispricing) model estimates. The column Diff. refers to the difference between baseline- and PCA-based decomposition approach extracted risk- and mispricing-components and the column t-stat tests if related return differences are statistically distinguishable from zero. The sample period for all strategies ends in December 2016. */**/*** indicate significance at the 10%/5%/1% level according to Newey and West (1987) corrected standard errors.

			Risk component				Mispricing component			
Momentum Definition	Reference	Start	Baseline	PCA	Diff.	t-stat.	Baseline	PCA	Diff.	t-stat.
Firm age	Zhang (2006)	12/1973	1.37***	1.40***	−0.03	−0.52	0.50***	0.46***	0.03	0.56
High volume stock mom.	Lee and Swaminathan (2000)	07/1972	0.97***	0.94***	0.03	0.28	0.82***	0.84***	−0.02	−0.15
Dyn. vola.-scaled	Daniel and Moskowitz (2016)	07/1972	0.99***	0.97***	0.02	0.45	0.30***	0.32***	−0.02	−0.45
Decile 10-1	Jegadeesh and Titman (1993)	07/1972	0.49**	0.45	0.04	0.38	0.85***	0.87***	−0.03	−0.28
Const. vola.-scaled	Barroso and Santa-Clara (2015)	07/1972	0.80***	0.79***	0.01	0.17	0.45***	0.46***	−0.01	−0.14
Const. semi-vola.-scaled	Wang and Yan (2021)	07/1972	0.74***	0.74***	0.00	0.01	0.44***	0.44***	0.00	0.00
Junk stock	Avramov et al. (2007)	12/1978	0.81***	0.99***	−0.19**	−2.32	0.48***	0.30***	0.18**	2.26
FF3 residual	Blitz et al. (2011)	07/1972	0.76***	0.76***	0.00	0.09	0.34***	0.34***	0.00	0.00
Off season mom.	Heston and Sadka (2008)	07/1972	1.07***	1.08***	−0.01	−0.14	−0.19***	−0.18***	−0.01	−0.25
Principal customer	Cohen and Frazzini (2008)	07/1977	0.84***	0.64**	−0.06***	−4.19	0.09***	0.04***	0.06***	3.85
Return seasonality	Heston and Sadka (2008)	07/1972	0.71***	0.71***	0.01	0.95	0.00	0.00***	0.00	−0.64
Long-term rev./mom.	Chan and Kot (2006)	07/1972	0.48**	0.50**	−0.02	−0.34	0.52***	0.51***	0.01	0.18
Value-mom.-profitability	Novy-Marx (2013)	07/1972	0.59***	0.65***	−0.06	−1.35	0.44***	0.38***	0.06	1.26
UMD (standard mom.)	Jegadeesh and Titman (1993)	07/1972	0.29*	0.28	0.01	0.20	0.50***	0.50***	−0.01	−0.12
Industry	Moskowitz and Grinblatt (1999)	07/1972	0.18	0.23	−0.05	−0.87	0.44***	0.39***	0.05	0.88
Customer industry	Menzly and Ozbas (2010)	02/1986	0.71***	0.84***	−0.13***	−5.78	0.05**	−0.06***	0.11***	4.86
Sharpe ratio	Rachev et al. (2007)	07/1972	0.26**	0.27*	−0.01	−0.20	0.40***	0.39***	0.01	0.25
Intermediate	Novy-Marx (2012)	07/1972	0.25	0.25	0.00	0.11	0.16***	0.16***	0.00	−0.04
Value-mom.	Novy-Marx (2013)	07/1972	0.40*	0.41*	−0.01	−0.27	0.47***	0.47***	0.00	−0.02
Supplier industry	Menzly and Ozbas (2010)	02/1986	0.32	0.50**	−0.18***	−5.32	0.04	−0.13***	0.16***	4.47
50/50 value/mom.	Asness et al. (2013)	07/1972	0.29***	0.22***	0.08***	3.29	0.23***	0.31***	−0.08***	−3.48
Industry-adj.	Cohen and Polk (1998)	07/1972	0.23**	0.26**	−0.03	−0.89	0.23***	0.19***	0.03	0.99
Factor mom.	Ehsani and Linnainmaa (2022)	07/1972	0.25***	0.24***	0.01	0.86	0.11***	0.12***	−0.01	−1.05
PC factor mom. 1-10	Ehsani and Linnainmaa (2022)	01/1973	0.15***	0.16***	0.00	−0.78	0.04***	0.04***	0.00	0.65
PC factor mom. 11-20	Ehsani and Linnainmaa (2022)	01/1973	0.12***	0.13***	−0.01***	−3.19	0.03***	0.02***	0.01***	3.12
PC factor mom. 21-30	Ehsani and Linnainmaa (2022)	01/1973	0.06**	0.07***	−0.01*	−1.66	0.02***	0.01***	0.01*	1.69
PC factor mom. 31-40	Ehsani and Linnainmaa (2022)	01/1973	0.09***	0.10***	−0.01**	−2.16	0.04***	0.03***	0.01**	2.09
PC factor mom. 41-47	Ehsani and Linnainmaa (2022)	01/1973	0.12***	0.14***	−0.02***	−4.46	0.02***	0.00***	0.02***	4.21

A.III IPCA factor decomposition

Recently, Kelly et al. (2019) developed the instrumented principal component analysis (IPCA) that simultaneously estimates latent risk factors and time-varying betas. If the relationship between observable characteristics - that instrument for the unobservable dynamic loadings - and expected returns is driven by compensation for exposure to latent risk factors, IPCA can identify the corresponding latent factors and betas.

Kelly et al. (2019) conclude that the few characteristics that enter the IPCA factors (among them momentum) help to explain assets' exposures to systematic risks, not because they represent anomalous compensation without risk. In their five-factor model specification, Factor 3 is 50% correlated with UMD momentum which offers a further robustness test for our return decomposition approach in two respects: First, IPCA Factor 3 mainly loads on the momentum characteristic and is specifically constructed to capture systematic risk, so a decomposition of its returns is expected to extract a large risk-component and a negligible mispricing-component. Second, IPCA Factor 3 also loads significantly on other characteristics than momentum, particularly on short-term reversal, turnover, cash flow-to-book, unexplained volume, capital turnover, and return on assets. A return decomposition is thus implicitly controlling for these loadings and demonstrates the robustness of our approach in presence of other predictors for the cross-section of stock returns.

To separately analyze the long and short leg of IPCA Factor 3, we calculate each leg's return based on the following first-order condition:

$$f_{k,t} = \left(\beta'_{k,i,t}\beta_{k,i,t}\right)^{-1} \beta'_{k,i,t} r_{i,t}. \tag{4.17}$$

In month t, f_k denotes the return of factor k. With factor loadings $\beta_{k,i}$ $(N \times K)$ and stock returns r_i $(N \times 1)$, this means that $\left(\beta'_{k,i,t}\beta_{k,i,t}\right)^{-1} \beta'_{k,i,t}$ is $K \times N$ and can be interpreted as factor weights, i.e., the weight on each asset return that adds up to the factor realization.[42] We allocate stocks into portfolio legs depending on the sign of their weights, so stocks with positive weights are assigned to enter the long leg and otherwise. For the period July 1972 to May 2014, IPCA Factor 3 generates an average return of 2.27% p.m. with an annualized Sharpe ratio of 1.48. This return is entirely generated by its long leg having an average return of 2.28% and a correlation of 0.87 with UMDs long

[42] We thank Seth Pruitt for suggesting the use of this condition to separately construct long and short factor returns and Kelly et al. (2019) for making their code and data publicly available.

leg. The short leg contributes only an insignificant 0.01% while being highly correlated (0.92) with UMDs short leg.

Our decomposition results for IPCA Factor 3 as robustness tests are easily summarized: As expected, the average return of 2.27% (t-statistic: 8.25) contains a 2.31% (t-statistic: 8.93) risk-component, an insignificant -0.02% mispricing-component, and a highly significant option-component of -0.02%. Controlling for other significant factor loadings, our decomposition approach accurately captures the risk-component to a great extent. Nevertheless, (risk) factors derived from IPCA could be biased from correlated errors in investor expectations and thus reflect fluctuations in collective mispricing. This is exactly what we observe when looking at each portfolio leg: Most of the long legs return of 2.28% stems from a highly significant risk-component of 2.95% but it also contains a highly significant -0.52% mispricing-component. The latter is, however, offset by a mispricing-component of -0.50% included in the short leg, thus neutralizing common mispricing in the overall IPCA Factor 3 return. In contrast to the previous section, the return series under investigation is now specifically constructed to reflect risk and our decomposition approach successfully detects the absence of common mispricing in IPCA derived factor returns.

A.IV Rolling 20-year averages of UMD components

Figure A.I shows the dynamics of risk-, mispricing-, and option-components in value-weighted long-short momentum portfolio returns (UMD) over time.

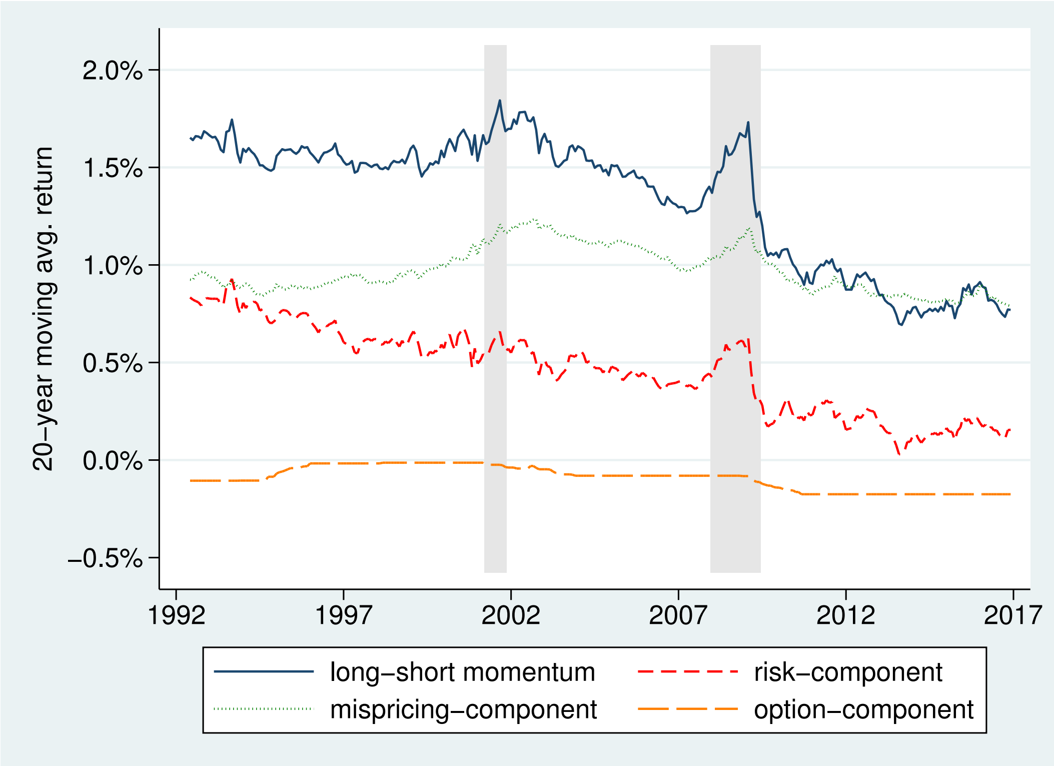

Fig. A.I. This figure shows rolling 20-year averages of monthly value-weighted long-short (decile 10-1) momentum portfolio returns decomposed into a risk-, mispricing-, and option-component. The gray shaded areas represent periods dated as recessions by the NBER. The sample period is July 1972 to December 2016.

We plot the 20-year moving averages of the components and find that momentum returns tend to decline in most recent decades. The 20-year average monthly return was 1.65% in June 1992 which reduces to only 0.77% in December 2016. Most of this decline in UMD returns over the last decades is attributable to the risk-component. In general, the magnitude of the mispricing-component exceeds the risk-component. The mispricing-component is relatively stable and increased from 0.92% to 1.19% until February 2009, whereas the risk-component dropped from an initial 0.83% to 0.62%. We observe a negative payoff for the option-component after times of momentum crashes when the market begins to rebound, reflecting the empirical findings in Daniel and Moskowitz (2016).

A.V Spanning regressions

In Section 4.4.3, we use 5 × 5 size-momentum portfolios as test portfolios to evaluate the performance of empirical asset pricing models. A huge body of literature emphasizes that these traditional models are not able to describe differences in expected returns for portfolios sorted on the momentum characteristic. Given this failure, our key question in Section 4.4.3 is: Do these models at least capture return differences among the risk-component of 5 × 5 size-momentum portfolios? This section presents results for a simple time-series spanning regression of the UMD risk-component on the market portfolio excess return (CAPM), risk factors proposed in Fama and French (2015) (FF5), and q-theory based factors proposed in Hou et al. (2015) (HXZ).

Table A.V: Spanning regressions for the UMD risk-component.

This table shows estimated coefficients for spanning regressions with the UMD risk-component as dependent variable. Our independent variables are the monthly returns for the value-weighted market portfolio in excess of the risk-free rate (CAPM), risk factors proposed in Fama and French (2015) (FF5), and q-theory based factors proposed in Hou et al. (2015) (HXZ). To obtain UMD risk-component returns, we use our baseline approach described in the main text (Section 4.3). Newey and West (1987) corrected t-statistics are given in parenthesis and */**/*** indicate significance at the 10%/5%/1% level. The last row reports R^2 adjusted for the degrees of freedom. The sample period is July 1972 to December 2016.

	CAPM	FF5	HXZ
Intercept	0.26*	0.34*	−0.09
	(1.81)	(1.87)	(−0.45)
MKTRF	0.05	0.01	0.07
	(0.78)	(0.23)	(1.29)
SMB		0.05	
		(0.52)	
HML		(−0.27)**	
		(−2.13)	
RMW		−0.12	
		(−0.66)	
CMA		0.22	
		(1.08)	
ME			0.24*
			(1.93)
I/A			0.05
			(0.31)
ROE			0.48***
			(3.38)
Adj. R^2	0.00	0.04	0.13

Estimated intercepts in Table A.V are statistically only weakly significant, implying that already the market factor (CAPM) is sufficient to describe the risk-component of UMD momentum returns.

A.VI A machine-learning approach (XGBoost) to identify macroeconomic and financial drivers of the risk-component of momentum returns

In this section, we extend the analysis presented in Section 4.4.5 and address a much broader question, which kind of risk is generally reflected by the risk-component of momentum returns?

To address this question, we combine a total of 126 U.S. macroeconomic variables (FRED-MD) provided by the Federal Reserve Bank of St. Louis (see McCracken and Ng (2016)) with 43 anomaly portfolios unrelated with momentum strategies proposed in Kozak et al. (2018) and Haddad et al. (2020).[43] Using this large panel as candidate proxies for macroeconomic and financial risks, we examine how well they describe the returns of the risk-component of UMD momentum. To investigate this relationship, we use machine learning, in particular, XGBoost, a gradient boosting decision tree algorithm. Introduced in Chen and Guestrin (2016), XGBoost and similar decision tree-based algorithms are considered best-in-class for small to medium tabular data at the moment and have many desirable features like invariance to variable scaling or robustness to outliers. The algorithm is able to uncover nonlinearities and interactions that are hard to be considered in linear regressions (see Bogousslavsky et al. (2024)).

For our purpose, we fit XGBoost to capture which of the 169 macroeconomic and financial variables ("features") covary with the returns of the risk-component of UMD momentum $r^{RISK}_{UMD,t}$. We follow a standard approach in machine learning and split the sample into a training (80%) and test (20%) subsample for evaluation. To find the optimal set of parameters (hypertuning), we perform grid search across a large range of possible combinations of hyperparameters (12,285 evaluations) using 4-fold cross-validation. This avoids overfitting and ensure the generalizability of our results. Performance metrics reveal the superiority of this machine learning approach. XGBoosts' out-of-sample R^2 is 23.15% compared to 9.86% using OLS and the RMSE reduces from 5.75 to 2.98. Using the full set of data (in-sample), the R^2 increases from 20.27% (OLS) to 88.09%. Overall, the model performance suggests that XGBoost yields a much better performance than OLS.[44]

[43] The macroeconomic variables are organized into eight categories: (1) output and income, (2) labor market, (3) consumption and orders, (4) orders and inventories, (5) money and credit, (6) interest rates and exchange rates, (7) prices, and (8) stock market. We exclude seven from the originally 50 anomaly portfolios because they are implementing momentum strategies. This avoids the circular argument of explaining momentum by itself.

[44] Differences between OLS and XGBoost metrics are all statistically highly significant using Welch's

To reveal the importance of the 126 macroeconomic and 43 financial variables for describing the risk-component of UMD returns, we use the SHapley Additive exPlanations (SHAP) method proposed in Lundberg and Lee (2017). Inspired by game theory, for each observation and feature, a SHAP value measures a weighted-average gain from adding a specific feature to all possible feature subsets.

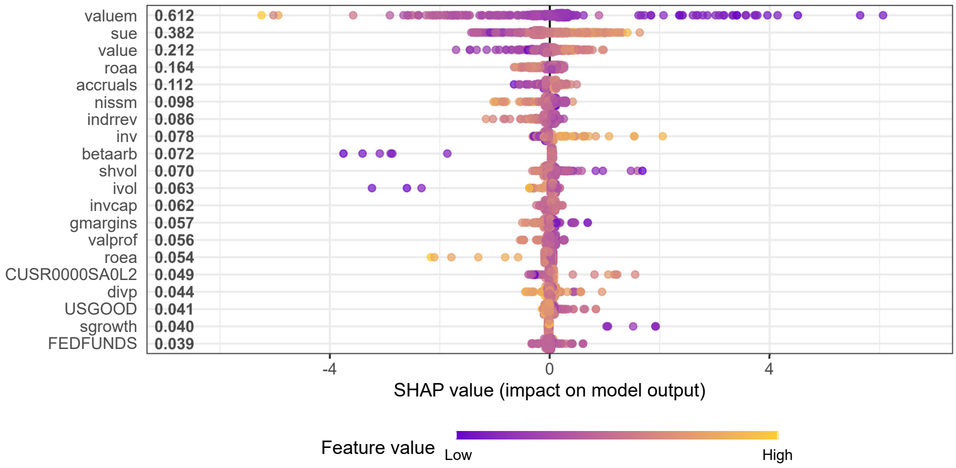

Fig. A.II. This figure ranks macroeconomic variables by their importance according to SHapley Additive exPlanations (SHAP) values (Lundberg and Lee (2017)). We only show the top 20 variables for brevity. We use a total of 126 macroeconomic variables (FRED-MD, see McCracken and Ng (2016)) and 43 financial anomaly portfolios unrelated with momentum strategies (Kozak et al. (2018) and Haddad et al. (2020)) as features.

Figure A.II ranks our twenty most important macroeconomic and financial variables according to SHAP and plots the distribution of SHAP values over all UMD return observations. Each point on the summary plot is a SHAP value for a feature and an instance (i.e., a single month of UMD's risk-component return). The color of each point represents the high (yellow) to low (blue) values of the feature. This provides insight into how the marginal contribution of a feature changes according to its value.

The most important determinant of UMD returns' risk-component is the monthly rebalanced value strategy "valuem" using quarterly updated book-to-market ratios proposed in Asness and Frazzini (2013). In line with Asness et al. (2013), we detect an inverse relation between "valuem" and momentum (risk-component) returns.[45] Notable, the complex relation between value and momentum is highlighted by the fact that traditional

two-sided t-test with bootstrapped standard errors.

[45]At the end of June of each year, "value" uses book equity from the previous fiscal year and market equity from December of the previous year and the portfolio constituents are compiled annually. Following Asness and Frazzini (2013), "valuem" calculates book-to-market ratios using the most up-to-date prices and book equity (appropriately lagged using quarterly accounting figures).

value strategy returns (i.e., using book equity from the previous fiscal year and market equity from December of the previous year, see Fama and French (1993)) is indeed positively covarying with momentum (risk-component) returns. For a better understanding of this relationship, Figure A.III shows the SHAP feature dependence between "value" (Fama and French (1993)) and "valuem" (Asness and Frazzini (2013)) for describing the risk-component of UMD momentum returns. The red line is a smooth curve fitted by LOESS (Cleveland (1979)).

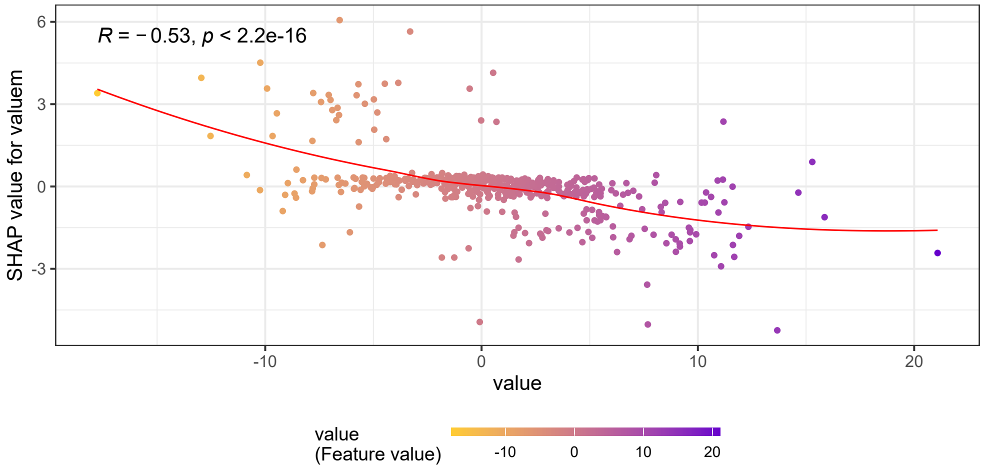

Fig. A.III. This figure shows the SHAP feature dependence for "value" (Fama and French (1993)) and "valuem" (Asness and Frazzini (2013)) for describing the risk-component of UMD momentum returns.

Together with the SHAP dependency plot shown in Figure A.III, our findings conclude that the risk-component of UMD momentum returns negatively covaries with "valuem" and positively covaries with "value", but low "value" returns give more rise for "valuem" to describe the risk-component of UMD.

Besides value, "sue" (standardized unexpected earnings, see Foster et al. (1984)) is ranked among the top three variables to determine UMD's risk-component and we observe that related returns positively covary. Generally, the influence of macroeconomic variables seems to be rather weak. Only three of the top 20 variables are related with macroeconomics: (1) CUSR0000SA0L2 (CPI: All items less shelter), (2) USGOOD (All Employees: Goods-Producing Industries), and (3) FEDFUNDS (Effective Federal Funds Rate). This is surprising, because the literature on momentum and its relation with macroeconomic risk so far emphasizes that most important sources seem to be the either the default spread (Asness et al. (2013)), industrial production growth (Liu and Zhang

(2008)), or one of the factors[46] proposed in Chen et al. (1986) (see Cooper et al. (2022)).

Our machine learning approach allows for multiple testing in the sense of Harvey and Liu (2016) and Feng et al. (2020) to address whether a given macroeconomic or financial variable adds explanatory power beyond what can be explained by the high-dimensional set of more than hundred of other variables. However, looking at the top 20 most important features of UMD's risk-component, no clear economic story emerges from these results. We do, however, clearly detect that value strategy returns are the most important determinant.

[46]These are (1) *MP*, the growth rate of industrial production, (2) *UI*, unexpected inflation, (3) *DEI*, change in expected inflation, (4) *UTS*, the global term premium, and (5) *UPR*, the U.S. default spread.

Chapter 5

Diversification Benefits of Luxury Watches and Day-of-the-week Effects in a Seven-Day Traded Market

This research is joint work with Klaus Röder (University of Regensburg). The paper is currently under review in the *Journal of Banking & Finance*. The journal ranking is *A* according to the VHB Publication Media Rating 2024.

Abstract

We show that luxury watches - particularly Rolex, Patek Philippe, and Audemars Piguet - yield significant diversification benefits when being added to well-diversified portfolios comprising stocks, bonds, and gold, and even outperform them on a risk-adjusted basis. All luxury watch returns exhibit remarkable low volatility most comparable with bonds while being uncorrelated with traditional asset classes. Evaluating day-of-the-week effects, this study also identifies that watch returns are generally lower on Sundays. This is likely because most sellers are professional dealers who do not edit their offers on Sundays, which is typically a day of rest.

Keywords: Empirical Asset Pricing · Alternative Investments · Collectibles · Luxury Watches

JEL classification: G10 · G11.

5.1 Introduction

Luxury watches as a financial investment have been overlooked by academics. This is surprising in two ways: First, risk and return properties of other collectibles are broadly studied in recent years, e.g., art (Korteweg et al. (2016)), diamonds (Auer and Schuhmacher (2013)), fine wine (Le Fur and Outreville (2019)), whiskey (Moroz and Pecchioli (2019)), and comics (Bocart et al. (2023)). Second, according to the latest available Deloitte Art & Finance Report 2023, the wealth of ultra-high-net-worth individuals associated with art and collectibles was already an astonishing $2.174 trillion in 2022, highlighting that an increasing number of people are willing to invest in these alternative investment classes.[1] The growth potential of the luxury collectibles market is evident by its auction sales surge over the last years. In May 2023 alone, luxury watch auctions in Geneva netted Phillips, Sotheby's, and Christies's over CHF 100 million in sales (see Deloitte (2024)). The report further states that client allocation of surveyed wealth managers to art and collectibles was 10.9% in 2023, where private banks reported an average of 8.6% allocation and family offices of 13.4%. Given these numbers, it is incomprehensible that the luxury watch market has received little academic attention. We fill this gap in the literature.

To the best of our knowledge, Köstlmeier and Röder (2025b) is the only study analyzing the global market for luxury watches. Their main focus, however, is on analyzing watch counterparts for well-known stock market anomalies through the lens of asset pricing. They find that size, reversal, max, and momentum form successful long-short strategies among luxury watches. Using a novel dataset in this study, we examine an important empirical question that arises to any investor when being confronted with new investment prospects: Do luxury watches provide additional diversification benefits beyond stocks, bonds, and gold, and if so, are there potential day-of-the-week effects that should be accounted for when buying or selling luxury watches?

This study provides several key contributions to the literature. First, analyzing daily and weekly returns of six luxury watch indices representing the brands Rolex, Patek Philippe, Audemars Piguet, Cartier, Omega, and Tudor from January 2017 to September 2024, we find that some of them generate quite large returns. The average annualized return of Rolex (6.94%), Patek Philippe (10.61%), and Audemars Piguet (10.81%) is close to the performance of U.S. stocks (9.28%). We observe that return volatility of luxury

[1] See Deloitte (2023). An ultra-high-net-worth individual is someone with a net worth of $30 million or more.

watches is remarkably low and just about one fifth of stock market volatility, thus quite the same as Treasury bills. Taken together, the annualized Sharpe ratios of Audemars Piguet (1.55), Patek Philippe (1.29), and Rolex (1.26) vastly exceeds the Sharpe ratio of the U.S. stock market (0.71) which is somehow surprising given its remarkable well performance in recent years.

Second, we analyze the diversification potential of luxury watches and the key question if benefits stem from an increase in portfolio returns, a reduction of risk, or both together. We use several approaches to address this question. In lack of an appropriate asset pricing model for luxury watches, we first examine correlations with traditional asset classes and find that correlations of luxury watches with stocks, bonds, and gold are essentially close to zero. Besides this static analysis, correlations may be different at other points in time so we use a dynamic conditional correlation (DCC-) GARCH model proposed in Bollerslev (1986) and Engle (2002) to evaluate time-varying correlations. We are not able to reject the null hypothesis of constant, time-invariant correlations for all six luxury watch indices, indicating that the observed low correlation prevails throughout our sample period. Applying conditional quantile correlations, we conclude that luxury watches are a reasonable hedge against plunges in stocks, bonds, and gold, but only a safe haven of weak nature in the meaning of Baur and Lucey (2010) and Baur and McDermott (2010).[2]

Third, we conduct mean-variance spanning tests as proposed in Huberman and Kandel (1987), Ferson and Keim (1993), and Kan and Zhou (2012) to explore the economic source of diversification benefits. Given their remarkable well return performance, Audemars Piguet, Patek Philippe, and Rolex improve portfolio returns of benchmark assets. In other words, they shift the resulting tangential portfolio comprising stocks, bonds, and gold towards higher risk-adjusted returns. This is, however, not the case for Omega, Cartier, and Tudor, as their returns are close to zero. However, all luxury watches significantly improve the minimum-variance portfolio, thus lowering the risk embodied in the portfolio of benchmark assets.[3] This is mainly because all luxury watch portfolios lie outside the efficient frontier formed by stocks, bonds, and gold. To illustrate the diversification benefit, an investor who targets an annual return of 5% would hold 70.61% bonds, 16.05% stocks, and 13.34% gold, achieving an annualized Sharpe ratio of 0.61. With no short sales allowed, the Sharpe ratio doubles to 1.23 when adding luxury watches because

[2]A hedge requires asset returns to be uncorrelated *on average*. A safe haven is defined similarly, but for times of *stress*.

[3]As a matter of fact, daily returns for all luxury watch indices are highly non-normal and with the exception of Cartier positively skewed, so deviations from the mean are mainly due to beneficial upwards movements.

return volatility almost halves to only 2.35% p.a. Last, spanning regressions reveal that estimated alphas are a large 5.28% p.a. in case of Rolex and exceed even 10% p.a. in case of Patek Philippe and Audemars Piguet despite controlling not only for contemporaneous, but also for up to four weeks lagged, returns of benchmark assets.

Fourth, as our main results demonstrate that investors benefit from a diversification potential of luxury watches, we address the question that arises when implementing the actual watch investment: When is the best time, in the sense of which day of the week, to buy them? This question when to invest is especially interesting for luxury watches since they can be continuously traded all seven days of the week.[4] Given that daily watch returns are positively skewed, have a very high excess kurtosis, and are generally non-normal distributed, we apply an ARMA-EGARCH(-M) model to analyze daily differences in the returns and volatilities of luxury watch portfolios. In general, we document that returns on Sundays are significantly lower than on other days. Similarly, a negative Sunday effect is also detected for the conditional variance of our luxury watch portfolios under study, however, it is less pronounced compared to the conditional mean. Our additional analysis suggests that the negative Sunday effect is caused by the predominance of professional dealers that mostly act as sellers of luxury watches on large peer-to-peer marketplaces. Given that Sunday is typically a day of rest in most western countries, they do not edit, resp., update their offers on that day to such an extend as on other days of the week.

The rest of the paper is organized as follows. Section 5.2 provides information and descriptive statistics for our data on luxury watches. Section 5.3 provides details and results for the diversification benefits of luxury watches. To begin with, Section 5.3.1 provides a performance analysis of luxury watches and evaluates potential diversification benefits by analyzing static correlations with stocks, bonds, and gold. Then, Section 5.3.2 extends these findings by studying related time-varying correlations with these traditional asset classes. What follows are tests for the diversification benefits of luxury watch watches in a mean-variance framework presented in Section 5.3.3. Section 5.3.4 analyzes the impact of transaction costs on our main results. Section 5.4 conducts tests on potential day-of-the-week effects for luxury watch returns and includes possible explanations for the underlying economic channels. Finally, Section 5.5 concludes.

[4] The two largest peer-to-peer platforms for trading luxury watches, WatchCharts Marketplace and Chrono24 list more than 914,000, resp. 564,000, watches for sale as of April 2025. More than 90,000 watches offered via WatchCharts Marketplace cost more than $10,000, which clearly qualifies them as luxury timepieces.

5.2 Data and descriptive statistics

The data set comprises daily price data for luxury watches from WatchCharts Analytics. Our sample period is 01/01/2017 to 09/30/2024 and all data is denominated in currency U.S.-$. WatchCharts provides price information for a broad range of watch indices at the brand level. They collect and analyzes millions of sales from across the web for constructing these indices, relying on three different sales categories: (i) dealer sales, i.e., watches sold by professionals, (ii) private sales, and (iii) auction sales. Based on their own corporate statements, most data stems from dealer sales and takes the form of ask-prices. Price information is retrieved from popular platforms and secondary marketplaces around the world, among them e.g., eBay, Reddit, Rolex-Forums, Omega-Forums, ManOnTime, Rakuten (Japan), or Carousell (Asia-Pacific). Each brand index tracks the performance for a specific brand, based on the 30 top watches from the brand.

To determine which watches are in an index, they estimate the market share (based on estimated total annual transaction value) for each watch that is eligible for inclusion in a particular index, and select the top watches based on this metric. Similarly, they also use market share to determine the weighting of each watch. The index performance is then calculated as a weighted average of the price performance of the set of watches. To reflect market trends over a long-term period, the set of watches and weights for each index is determined on a calendar year basis: Every January 1st, they update the set of watches and weights for the new year based on the top watches matching the inclusion criteria in the prior year.

The luxury watch market is dominated by few Swiss manufactures. According to Morgan Stanley's annual watch report (see Müller (2024)), the top five leading Swiss manufactures in terms of worldwide retail market share 2023 are Rolex (30.3%, including their brand Tudor), Cartier (7.5%), Omega (7.5%), Patek Philippe (5.6%), and Audemars Piguet (4.9%). Reflecting their economic relevance, our sample comprises related indices for these six brands covering a combined market share of 55.8%. The dominance of these brands among luxury watches is also documented in Köstlmeier and Röder (2025b) and in line with reports on sales activities on Chrono24 (see Chrono24 (2022) and Deloitte (2024)), the worlds largest peer-to-peer platform for trading luxury watches.[5]

WatchCharts does not disclose historic constituents for each index, but as of September

[5] Based on own corporate statements, Chrono24 is the worldwide leading platform for luxury watches since 2003 with more than 25,000 trustworthy sellers, 9 mio. active users each month, and more than 548,000 offered watches as of November 2024.

30, 2024, top holdings in each index are the watches Rolex Datejust (ref. no. 126334, 8.7%), Audemars Piguet Royal Oak (reference no. 15202ST.OO.1240ST.01, 6.3%), Patek Philippe Nautilus (reference no. 5712/1A, 8.3%), Cartier Santos (reference no. WSSA0030, 11.1%), Omega Speedmaster Silver Snoopy (reference no. 310.32.42.50.02.001,9.7%), and Tudor Black Bay (reference no. 79360N, 12.4%). Compared to prominent stock indices like e.g., S&P500, where the top ten constituents' amount for more than 30% of aggregated market capitalization nowadays, the moderate weights of each top holding in our analyzed watch indices indicate that they are well diversified portfolios, combining a total of 30 luxury watches for each brand.

To examine diversification benefits of luxury watches from the perspective of a U.S. investor, we compare them with distinct asset classes covering a broad range of investment possibilities. The equity market is represented by the U.S. value-weighted market return provided by K. French. The index of the S&P U.S. Treasury Bond Index, which is a market-value weighted index, measures the performance of the U.S. Treasury Bond market. Finally, we include the price of Gold Bullion per Troy ounce (LBMA p.m. fixing). Any investing in these benchmark portfolios displays a well-diversified portfolio. Weekly closing prices are retrieved from LSEG Refinitiv/Datastream and also reach from 01/01/2017 to 09/30/2024. We consider total return indices denoted in currency U.S.-$.

Table 5.1: Daily watch returns.

This table shows the descriptive statistics of daily log returns (in % and denoted in U.S.-$) for the six luxury watch indices under study until 09/30/2024. */**/*** indicates that coefficients are significantly different from zero at a level of 10%/5%/1%. We use Newey and West (1987) corrected standard errors for testing the mean return with a lag equal to the integer part of $4(N/100)^{2/9}$.

Watches	Start	Mean	SD	Skew.	Exc. kurt.	Min.	25th	Median	75th	Max.	Jarque-Bera	ADF (const.)	N
Audemars Piguet	01/02/2019	0.0296**	0.18	2.30	20.30	−1.05	−0.04	0.00	0.08	2.14	37,895.36***	−22.93***	2099
Cartier	01/02/2020	0.0002	0.18	−0.56	11.97	−1.92	−0.09	0.00	0.09	1.28	10,441.98***	−31.14***	1734
Omega	01/02/2017	0.0081*	0.25	1.03	19.04	−1.62	−0.08	0.00	0.09	3.21	43,234.58***	−40.49***	2829
Patek Philippe	01/02/2019	0.0291**	0.22	5.14	89.90	−0.96	−0.04	0.00	0.08	4.56	716,037.40***	−23.86***	2099
Rolex	01/02/2017	0.0191***	0.14	0.60	4.19	−0.70	−0.05	0.01	0.08	1.13	2,238.93***	−28.89***	2829
Tudor	01/02/2017	0.0003	0.18	0.99	17.44	−1.51	−0.06	0.00	0.05	1.77	36,320.82***	−37.60***	2829

Table 5.1 presents summary statistics for the time-series of daily log returns of the six luxury watch indices. Because of data availability restrictions, Audemars Piguet and Patek Philippe start in 2019 and Cartier in 2020. Daily returns for all watch brand indices are positive, only slightly volatile, and with the exception of Cartier, positively skewed. These distributional properties of luxury watch returns may not be disadvantageous to investors because they imply high outliers. The Jarque-Bera test staistic indicates highly non-normal distributions for all six luxury watch indices under study and the augmented Dickey-Fuller (ADF) test statistic (including a constant) indicates stationary for all return series at a significance level of 1%. The portfolio comprising Rolex watches yields a highly significant 1.91 bps per day while having the lowest volatility of only 0.14%. Audemars Piguet and Patek Philippe watches show the highest average daily returns with 2.96 bps and 2.91 bps, whereas Cartier generates only 0.02 bps on average. Annualized, this amounts to an average return of 10.81% (Audemars Piguet), 10.61% (Patek Philippe), and 6.94% (Rolex).[6] Because of the remarkable low volatility of luxury watches, annualized Sharpe ratios are an astonishing 2.16 (Audemars Piguet), 1.75 (Patek Philippe), and 1.48 for the Rolex index.[7] Nevertheless, annualized Sharpe ratios are a negative -0.89 for Cartier and -0.77 for Tudor. Volatility and skewness of Patek Philippe and Audemars Piguet are among the highest compared to other brands. 25% of all returns for Patek Philipp exceed 8 bps and we observe a maximum return of 4.56% on a single day. For that reason, it is important for investor to time when to trade luxury watches, especially at which day of the week. In contrast to equities which can only be traded on five business days, luxury watches can be traded throughout each day of the week. Similarly, cryptocurrencies are also continuously traded and related studies show that both returns and related volatilities behave differently on weekends (see Dorfleitner and Lung (2018), Qadan et al. (2022)). We will analyze these potential day-of-the-week effects in Section 5.4.

5.3 Diversification potential of luxury watches

5.3.1 Performance analysis

Besides looking at luxury watches as isolated investments, we analyze in this section how they behave in a portfolio context together with stocks, bonds, and gold. Based on the

[6]Luxury watches are traded on peer-to-peer platforms like Chrono24, so the number of trading days is 365 per year.

[7]We calculate the Sharpe ratio using discrete returns. The risk-free interest rate is the one-month Treasury bill rate provided by K. French.

findings in Dorfleitner (2003) we use discrete returns which are preferred when examining primarily portfolio-related aspects of returns series. To illustrate the performance of these assets, Figure 5.1 demonstrates the cumulative payoff of a $1.00 investment in U.S. stocks (tight dotted), bonds (dotted), gold (dashed), and the portfolio comprising 30 Rolex watches (solid line).

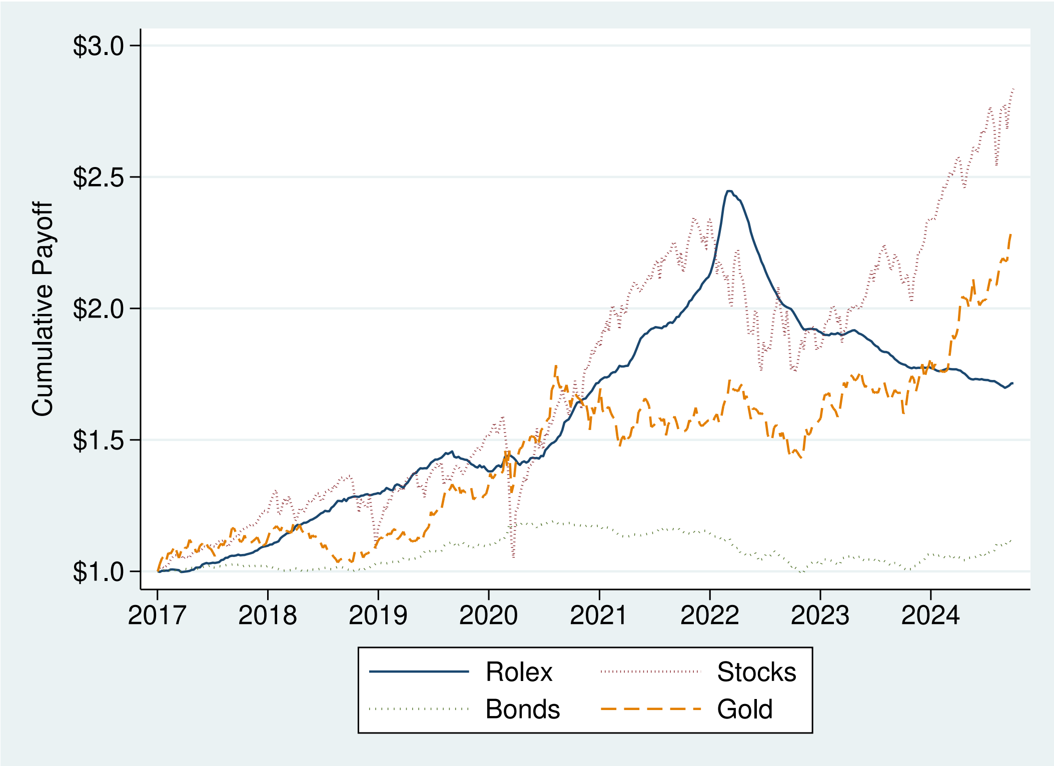

Fig. 5.1. Cumulative payoff of U.S. stocks (tight dotted), U.S. bonds (dotted) as represented by the S&P U.S. Treasury Bond Index, gold (dashed), and Rolex watches (solid line) of a $1.00 investment over the sample period until 09/30/2024.

For the sake of brevity and because of longer time-series data availability, we choose Rolex mostly because of their higher economic importance compared to other watch brands. Charts for other watch brand portfolios are quite similar, given that the correlation of weekly Rolex returns with other luxury watches is within the range from 0.24 (Tudor) to 0.55 (Patek Philippe). The initial investment of $1 grows to $2.84 until 09/30/2024 when investing in stocks, demonstrating that diversified stock investments reward large payoffs. The terminal value of gold is a remarkably high $2.30, outperforming the investment in Rolex watches ($1.71) and bonds ($1.11). The luxury watch indices under study generally show an upward movement in index values until the first half year 2022.[8] The maximum all time high of the Rolex investment was in 03/07/2022 with a payoff of $2.45 for each dollar investment at the beginning of 2017 thus outperforming stocks with a payoff of $2.02 at that time. After that, however, we notice a remarkable decrease in index values until the end of our sample period. For that reason, we split our overall sample period into two consecutive sub-periods. The first sub-period starting 01/02/2017 and ending 03/07/2022 (Rolex peak) displays a large growth in luxury watch investments, while the second sub-period starting 03/14/2022 until 09/30/2024 is characterized by a strong downwards movement in luxury watch index values.

Table 5.2 reports descriptive statistics on weekly returns of the six luxury watch indices under study and the three benchmark assets, respectively. Most of the characteristics for continuous daily returns shown in Table 5.1 also hold for discrete weekly returns. For the full sample period, Patek Philippe and Audemars Piguet watch portfolios earn 0.21% per week, only outperformed by the 0.29% return generated by the stock market. What stands out in this table is the remarkable low standard deviation of watch returns. Patek Philippe exhibits the highest volatility with 0.90% and its minimum return is a moderate -2.03%. Overall, these results suggest that the volatility of watches is more comparable to U.S. Treasuries (0.58%), than stocks or even gold. In consequence of the low volatility, luxury watches of Audemars Piguet, Patek Philippe, and Rolex have a much higher Sharpe ratio than stocks. Together with the positive skewness of weekly returns and a negligible correlation with all benchmark assets, investments in luxury watches seems to have a beneficial diversification potential. The Covid-19 related stock market turmoil in March 2020 is a good illustration for that potential: While the U.S. stock market (gold) accumulated a loss of -27.78% (-4.64%) during the three weeks from 03/02/2020 to 03/23/2020, four out of six luxury watch portfolios generated positive returns as high as 1.32% (Tudor), only outperformed by bonds (2.32%). However, the worst luxury watch performance for that period was a moderate loss of -0.45% for Rolex watches.

Splitting the sample period into two sub-periods shows that different market phases have a tremendous impact on the characteristics of luxury watch returns. Panel B in Table 5.2 shows descriptive statistics for the bullish sub-period until 03/07/2022. Similar with our daily return analysis, we notice a discrepancy in the performance among brands. While Audemars Piguet, Patek Philippe, and Rolex earn a large return of 0.64%, resp. 0.66% and 0.33%, the portfolios comprising watches of Cartier, Omega, and Tudor only generate

[8]The only exception is Tudor, having its peak performance already in 06/07/2021.

Table 5.2: Weekly discrete returns of luxury watches and benchmark assets.

This table shows descriptive statistics of weekly discrete returns (in % and denoted in U.S.-$) for six luxury watch indices under study and benchmark assets. Each luxury watch index comprises a portfolio of 30 watches of the related brand and their returns are weighted by market shares. The benchmark assets are stocks (value-weighted U.S. market return), bonds (S&P U.S. Treasury Bond Index), and gold (LBMA p.m. fixing). The Sharpe ratio is expressed annualized. N denotes the number of observations included in each period. Correlation refers to the Pearson correlation. The sample period begins 01/02/2017 for Rolex, Tudor, and Omega, 01/02/2019 for Audemars Piguet and Patek Philippe, and 01/02/2020 for Cartier and ends 09/30/2024.

						Sharpe	Correlation			
	Mean	SD	Min.	Max.	Skew.	ratio	Stocks	Bonds	Gold	N
Panel A: Overall period, 01/02/2017 to 09/30/2024										
Audemars Piguet	0.21	0.76	−1.93	2.96	0.76	1.55	−0.04	−0.01	0.03	299
Cartier	0.00	0.46	−2.03	2.26	0.51	−0.63	0.01	−0.02	−0.01	247
Omega	0.06	0.57	−2.13	2.78	0.41	0.26	0.04	0.05	0.00	404
Patek Philippe	0.21	0.90	−2.03	4.76	1.50	1.29	0.02	−0.10	0.06	299
Rolex	0.14	0.54	−1.54	2.56	0.56	1.26	−0.05	0.01	0.06	404
Tudor	0.00	0.52	−1.73	2.63	0.79	−0.47	−0.05	0.07	0.00	404
Stocks	0.29	2.59	−13.80	16.82	−0.31	0.71	1.00	−0.08	0.15	404
Bonds	0.03	0.58	−1.94	3.14	0.61	−0.13	−0.08	1.00	0.37	404
Gold	0.22	1.89	−11.05	6.66	−0.33	0.71	0.15	0.37	1.00	404
Panel B: First sub-period, 01/02/2017 to 03/07/2022										
Audemars Piguet	0.64	0.70	−1.26	2.96	0.62	6.33	−0.12	−0.08	0.02	165
Cartier	0.07	0.51	−2.03	2.26	0.67	0.90	−0.04	−0.06	−0.03	113
Omega	0.14	0.61	−2.13	2.78	0.34	1.37	0.04	−0.01	0.00	270
Patek Philippe	0.66	0.85	−0.65	4.76	2.23	5.41	−0.01	−0.21	0.07	165
Rolex	0.33	0.51	−1.40	2.56	0.71	4.39	−0.10	−0.06	0.04	270
Tudor	0.10	0.55	−1.73	2.63	0.68	1.08	−0.08	0.03	−0.03	270
Stocks	0.30	2.63	−13.80	16.82	−0.23	0.76	1.00	−0.32	0.11	270
Bonds	0.04	0.46	−1.10	3.14	1.17	0.40	−0.32	1.00	0.39	270
Gold	0.22	1.94	−11.05	6.66	−0.48	0.75	0.11	0.39	1.00	270
Panel C: Second sub-period, 03/14/2022 to 09/30/2024										
Audemars Piguet	−0.32	0.42	−1.93	1.45	0.37	−6.72	0.07	−0.03	0.04	134
Cartier	−0.05	0.40	−1.11	0.98	0.00	−2.34	0.08	0.00	0.01	134
Omega	−0.10	0.43	−1.34	0.82	−0.22	−2.95	0.04	0.16	0.01	134
Patek Philippe	−0.35	0.58	−2.03	2.37	0.66	−5.39	0.04	−0.11	0.03	134
Rolex	−0.26	0.36	−1.54	0.48	−0.79	−7.34	0.07	0.07	0.16	134
Tudor	−0.20	0.38	−1.33	1.07	0.29	−5.22	0.05	0.12	0.11	134
Stocks	0.28	2.51	−9.30	7.55	−0.50	0.59	1.00	0.22	0.22	134
Bonds	−0.01	0.76	−1.94	2.47	0.36	−0.79	0.22	1.00	0.37	134
Gold	0.23	1.80	−4.77	5.36	0.06	0.61	0.22	0.37	1.00	134

returns that are less than 0.14%. Putting things into perspective, all luxury watch returns exceed the 0.04% return of bonds, although their volatility is very close to them. The Sharpe ratio of all luxury watches exceeds the 0.76 Sharpe ratio of stocks, and no return is less than -2.13, while the stock market lost -13.80% from 03/09/2020 to 03/16/2020 because of Covid-19 related stock market turmoils. In contrast, average luxury watch returns are negative for the second sub-period starting in 03/14/2022. The most striking result in Panel C is that the already low volatility of luxury watches is even less during the bear market. The standard deviation of Patek Philipp returns decreases from 0.85% in the first sub-period to 0.58% from 03/14/2022 and all volatilities are less than the observed volatility of 0.76% for bonds. Again, this is in stark contrast to stocks, where we typically observe in increase in volatility during downwards market phases (Nelson (1991), Glosten et al. (1993)).

5.3.2 Dynamic correlation analysis

Although we find very low correlation values close to zero for all luxury watch brand indices, the observed low correlations may be different at other points in time. For that reason, we estimate time-varying conditional correlations between luxury watch returns and stocks, bonds, and gold by using the multivariate GARCH model presented in Bollerslev (1986) and Engle (2002). According to this dynamic conditional correlation GARCH model (DCC-GARCH), a zero-mean vector r_t of k asset returns has a conditional variance-covariance matrix $H_t = \mathbb{E}_{t-1}\left(r_t r_t'\right)$ which can be represented as $H_t = D_t \Gamma_t D_t$, with $D_t = \text{diag}\sqrt{h_t}$ as a stochastic diagonal matrix holding the conditional standard deviations of r_t on its main diagonal and with Γ_t as a time-varying correlation matrix. In this model, $r_t = D_t \epsilon_t$ holds with the assumption $\epsilon_t \sim \mathcal{N}(0,1)$. The volatility series D_t is assumed to be represented by a GARCH(1,1) process such that

$$h_{i,t} = \omega_i + \alpha_i r_{i,t-1}^2 + \beta_i h_{i,t-1}, \tag{5.1}$$

for $i = 1, 2, ..., k$, with $\omega_i > 0$, $\alpha_i \geq 0$, $\beta_i \geq 0$, and $\alpha_i + \beta_i < 1$.

We estimate the dynamic dependence of the correlation matrix Γ_t with a DCC(1,1) model as proposed by Engle (2002):

$$\begin{aligned} \Gamma_t &= \text{diag}\left(q_{11,t}^{-1/2}, \ldots, q_{kk,t}^{-1/2}\right) Q_t \text{diag}\left(q_{11,t}^{-1/2}, \ldots, q_{kk,t}^{-1/2}\right), \\ Q_t &= (1 - \theta_1 - \theta_2)\bar{Q} + \theta_1\left(\epsilon t - 1\epsilon_{t-1}'\right) + \theta_2 Q_{t-1}, \end{aligned} \tag{5.2}$$

where Q_t is a $k \times k$-dimensional positive-definite matrix representing the conditional standardized residuals' variance–covariance matrix with $q_{ii,t}$ as its diagonal elements and $\bar{Q}$ its unconditional counterpart. Our focus of interest is the first row of Γ_t holding the time-varying conditional correlations between luxury watch brand portfolio returns and the returns for stocks, bonds, and gold. We estimate the DCC(1,1)-GARCH(1,1) model separately for each of our six luxury watch index returns together with stocks, bonds, and gold, but only report estimated coefficients of luxury watches for brevity.

Table 5.3: Time-varying correlation analysis.

This table shows quasi maximum likelihood estimation results for luxury watch portfolio returns of a DCC(1,1)-GARCH(1,1) model as specified in the main text. T-statistics based on White (1982) corrected standard errors are provided in parenthesis. */**/*** indicates that coefficients are significantly different from zero at a level of 10%/5%/1%. The χ^2 test by Engle and Sheppard (2001) examines the null hypothesis of constant conditional correlation. Our model includes time-series of weekly returns for luxury watch brand indices, value-weighted U.S. stock market returns, returns of the S&P U.S. Treasury Bond Index, and the returns of gold (LBMA p.m. fixing). For brevity, we only report coefficient estimates for luxury watch portfolio returns. The sample period begins 01/02/2017 for Rolex, Tudor, and Omega, 01/02/2019 for Audemars Piguet and Patek Philippe, and 01/02/2020 for Cartier and ends 09/30/2024.

	Rolex	P. Philippe	A. Piguet	Cartier	Omega	Tudor
GARCH(1,1) parameters						
ω	0.0011	0.1035	0.0000	0.0144	0.0056	0.0263*
	(0.7240)	(0.4377)	(0.0001)	(0.8315)	(1.1663)	(1.8929)
α	0.0858**	0.3052	0.0345	0.0643	0.0692**	0.1635***
	(2.3538)	(0.3232)	(1.4946)	(1.2633)	(2.2219)	(2.6436)
β	0.9101***	0.4764	0.9610***	0.8656***	0.9160***	0.7386***
	(22.5250)	(0.4042)	(51.3701)	(8.0845)	(28.8653)	(8.3698)
DCC(1,1) parameters						
θ_1	0.0281***	0.0436*	0.0384*	0.0553**	0.0352***	0.0262**
	(2.8345)	(1.7173)	(1.8379)	(1.9658)	(3.2626)	(1.9668)
θ_2	0.9213***	0.8973***	0.9025***	0.8283***	0.9083***	0.9334***
	(31.4791)	(14.4527)	(19.0289)	(14.5453)	(35.0571)	(25.7818)
Diagnostics						
N	404	299	299	247	404	404
χ^2-test	0.0031	1.4774	0.0139	0.0522	0.4045	0.6295
p-value	0.9985	0.4777	0.9931	0.9742	0.8169	0.7300
Log-likelihood	-2132.72	-1817.08	-1729.47	-1468.15	-2293.08	-2240.28

Table 5.3 shows quasi maximum likelihood estimation results for our model parameters and t-statistics based on White (1982) corrected standard errors in parenthesis. To begin with, the sums of estimated values for α and β are mostly close to one meaning that

the corresponding GARCH processes show a high degree of persistence. In line with the literature (e.g., Bates (2019) or Bollerslev and Todorov (2023)), the (unreported) coefficients α for stock market returns are within the range $[0.3449; 0.3959]$ and the corresponding decay estimates β are within $[0.5413; 0.9101]$, indicating that stock return volatility seems to be more sensitive to past volatility than past news shocks.[9] We observe highly significant estimates for β in all luxury watch indices with the exception of Patek Philippe. Estimated coefficients exceed 0.90 in case of Rolex, Audemars Piguet, and Omega, while being slightly less emphasized for Cartier (0.8656) and Tudor (0.7386). More interesting, the estimated DCC parameters θ_1 and θ_2 for luxury watch returns are statistically significant and imply persistent correlation. Engle and Sheppard (2001) suggest a χ^2 test for the null hypothesis that correlations Γ_t are constant over time. This test is not able to reject the null hypothesis for all luxury watch indices, indicating that our previous findings of low correlations between luxury watch returns and stocks, bonds, and gold are indeed constant and time invariant.

Derived from our model estimation results, Figure 5.2 displays the time-varying conditional correlations of luxury watch portfolio returns for the brands Rolex, Patek Philippe, and Audemars Piguet with value-weighted U.S. stock market portfolio returns.[10] First, it confirms our previous findings of low correlations close to zero during our sample period. Second, we notice that dynamic correlations fluctuate at a high degree and that no obvious structural breaks, separating regimes of different correlations, exist. In case of Rolex, the minimum cond. correlation is -0.27 in 01/31/2022 and a maximum of 0.13 in 08/12/2024. Their average correlation is only -0.03 with a standard deviation of 0.06. Interestingly, the correlation during Covid-19 related stock market turmoils in 03/16/2020 is a negative -0.11 for Rolex, resp. -0.27 for Audemars Piguet but a positive 0.16 for Patek Philippe.

After having estimated dynamic conditional correlations, the question arises whether luxury watches can act as a hedge or a safe haven against plunges in the stock market. According to Baur and Lucey (2010) and Baur and McDermott (2010), a hedge requires asset returns to be uncorrelated or negatively correlated *on average*. A safe haven is defined similarly, but for times of *stress*. Using the dynamic conditional correlations, we calculate both their average value over the entire sample and their average value conditional on crashes of the stock, bond, and gold market, defined as the according 10% and 5% quantiles of the most negative stock, bond, and gold returns during our sample

[9] Coefficient estimates for bonds are quite similar. However, α is statistically insignificant and close to zero for gold.

[10] We omit results for the brands Cartier, Omega, and Tudor for better visualization. Obtained correlations are, however, quite similar.

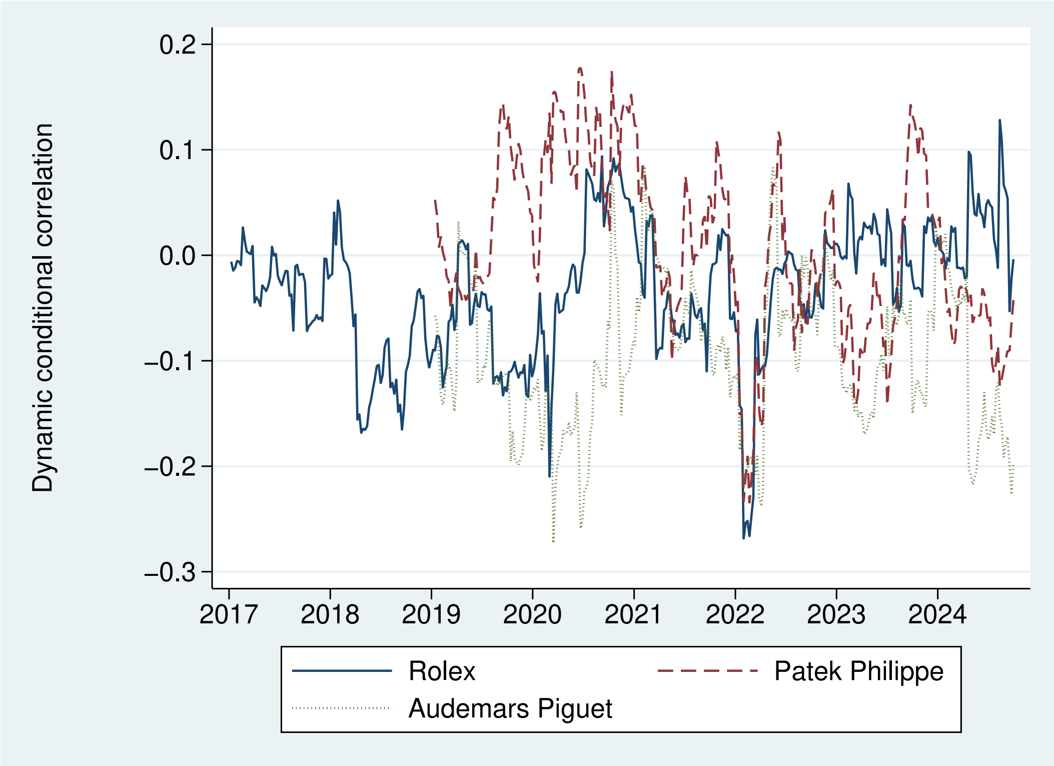

Fig. 5.2. This figure illustrates the conditional dynamic correlation of luxury watch portfolio returns of the brands Rolex, Patek Philippe, and Audemars Piguet with value-weighted U.S. stock market portfolio returns. The time-varying correlations are estimated using a DCC(1,1)-GARCH(1,1) model as specified in the main text.

period. These correlations are shown in Table 5.4.

In case of the stock market, a period of stress is indicated when weekly returns are below -2.57% (10% quantile), resp. -4.39% (5% quantile). We observe that all average correlations shown in Panel A are close to zero. Based on conventional statistical levels, none of the correlations is significant as indicated by the relatively high standard deviation of dynamic conditional correlations. In addition, average correlations according to the 10% quantiles (Panel B) or even 5% quantiles (Panel C) of the most negative stock, bonds, and gold returns are nearly identical to their unconditional averages. Taken together, the results in this section lead to the conclusion that luxury watches are indeed a hedge against the stock, bond, and gold market because on average, those returns are unrelated with each other. In times of market stress as indicated by the 5% quantiles of these benchmark assets' returns, luxury watches are at best only a safe haven of very weak nature, because their returns remain uncorrelated but don't seem to generate positive returns if stocks, bonds, or gold exhibit strong negative returns.

Table 5.4: Average and quantile correlations.

This table shows average values (Panel A) of the conditional correlations between luxury watch brand portfolio returns and stocks (value-weighted U.S. market return), bonds (S&P U.S. Treasury Bond Index returns), and gold (LBMA p.m. fixing) returns. Panel B (C) reports the average conditional correlations for the 10% (5%) quantile of the components' return distributions. Standard deviations of the correlations are provided in parenthesis.

	Rolex	P. Philippe	A. Piguet	Cartier	Omega	Tudor
Panel A: Average correlations						
Stocks	−0.03	0.01	−0.10	−0.04	0.04	−0.02
	(0.06)	(0.08)	(0.07)	(0.09)	(0.08)	(0.07)
Bonds	−0.04	−0.10	0.08	−0.05	0.03	0.06
	(0.07)	(0.08)	(0.07)	(0.08)	(0.08)	(0.06)
Gold	0.02	0.06	0.07	−0.05	−0.00	0.03
	(0.08)	(0.09)	(0.06)	(0.11)	(0.08)	(0.06)
Panel B: 10% quantile						
Stocks	−0.05	0.02	−0.09	−0.04	0.02	−0.05
	(0.07)	(0.10)	(0.08)	(0.09)	(0.09)	(0.07)
Bonds	−0.05	−0.11	0.08	−0.04	0.02	0.07
	(0.08)	(0.09)	(0.07)	(0.07)	(0.08)	(0.05)
Gold	0.04	0.07	0.09	−0.06	−0.00	0.05
	(0.08)	(0.08)	(0.05)	(0.10)	(0.07)	(0.06)
Panel C: 5% quantile						
Stocks	−0.07	0.02	−0.12	0.00	0.01	−0.07
	(0.07)	(0.12)	(0.09)	(0.07)	(0.08)	(0.09)
Bonds	−0.05	−0.10	0.07	−0.05	0.03	0.06
	(0.08)	(0.10)	(0.08)	(0.07)	(0.07)	(0.04)
Gold	0.05	0.07	0.08	−0.05	−0.01	0.06
	(0.08)	(0.09)	(0.06)	(0.11)	(0.07)	(0.07)

5.3.3 Mean-variance efficiency

Overview

The previous results indicate that returns of luxury watches do not seem to covary with the returns of stocks, bonds, and gold, thus enabling diversification potential. Besides returns, investors care about associated risk in the meaning of Markowitz (1952), so we turn to assess diversification benefits in a mean-variance framework in this section. We then conduct mean-variance spanning tests which check whether adding a given luxury watch to a well-diversified benchmark portfolio comprising stocks, bonds, and gold, improves the set of investment opportunities. In doing so, we address an important question: Do diversification benefits of luxury watches stem from an increase in portfolio returns, or from a reduction of risk?

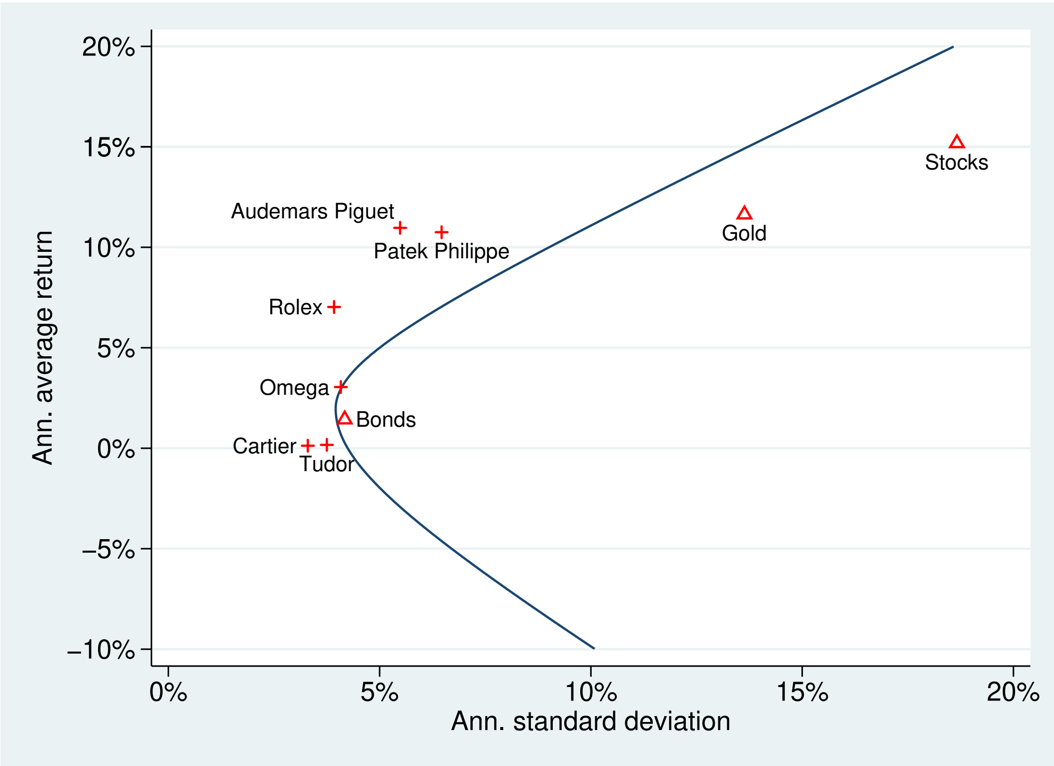

Fig. 5.3. This figure plots average returns and standard deviations of the value-weighted U.S. stock market portfolio, the S&P U.S. Treasury Bond Index, gold, and our six luxury watch portfolios. The blue line represents the ex post mean-variance efficient frontier from combining stocks, bonds, and gold. All numbers are in percentage points and annualized. The sample period is July 2017 to September 2024.

To obtain a better overview, Figure 5.3 plots the ex post opportunity set available to the investor from combining the U.S. value-weighted stock market portfolio, the S&P U.S.

Treasury Bond Index, and gold, for the period July 2017 to September 2024. We also add the sample return and standard deviation of the six luxury watch portfolios in the figure. Stocks had the highest average return of 15.18% p.a. among the benchmark assets, followed by gold (11.63%) and bonds (1.43%). What stands out is that all luxury watch portfolios lie outside the efficient frontier formed by our benchmark assets. This implies that an investor may be able to expand the opportunity set reliably by introducing some luxury watches into the portfolio. To begin with our analysis of diversification benefits in a mean-variance framework in more detail, we first take a look at portfolio weights and the maximum ex post mean-variance efficient (MVE) Sharpe ratios that can be achieved by combining various luxury watch portfolios with stocks, bonds, and gold to form the tangency portfolio.

Table 5.2 emphasizes that Rolex, Audemars Piguet, and Patek Philippe earn remarkably high returns that are uncorrelated with our benchmark assets. Cartier, Omega, and Tudor may not be as interesting in terms of generating high returns, but their low correlation with stocks, bonds, and gold may make them useful for diversification, too. For reference, the ex post MVE weights for maximizing the Sharpe ratio of our three benchmark assets are 30.90% for stocks, 28.62% for bonds, and 40.48% for gold.[11] This combination achieves an annualized Sharpe ratio of 1.11. Ex post MVE weights when adding a single luxury watch portfolio to stocks, bonds, and gold, are remarkably large, ranging from 12.77% (Tudor) to 77.74% (Rolex). Tudor's average daily return is only 0.03 bps, so any investment clearly is not able to boost performance but enhance risk-adjusted returns because of low volatility. In case of adding Rolex to our three benchmark assets, the ex post MVE weights for stocks, bonds, and gold drastically plunge to 7.60%, 7.77%, and 6.89%, whereas the Sharpe ratio increases from 1.11 (without Rolex) to 2.11 (with Rolex).

What are the optimal portfolio weights for other objectives?[12] An investor who targets an annual return of 5% using our benchmark assets would hold 70.61% bonds, 16.05% stocks, and 13.34% gold. The resulting portfolio return volatility is 5.00% p.a. and it achieves a Sharpe ratio of 0.61. With no short sales allowed, the Sharpe ratio doubles to 1.23 when adding luxury watches because volatility sharply decreases to only 2.35%. In this scenario, our benchmark assets comprise 22.64% of portfolio weights and most assets are invested in Rolex watches (26.88%). Even imposing a transaction cost constraint

[11] The large weight for gold is in line with Estrada (2016) who recommend a 36% weight if the original portfolio consists of 60% stocks and 40% bonds. However, based on traditional optimization asset allocation methods, Jaffe (1989) and Lucey et al. (2006) suggest a more moderate weight of 4%-10% for gold.

[12] We assume long-only positions with single asset weights up to 100% in this section.

amounting to 1% results in the mean-variance efficient portfolio to hold 20.40% luxury watches (mainly Rolex, Cartier, and Tudor), while comprising 20.60% stocks, 50.60% bonds, and 8.40% gold.

To put these investment allocations into perspective, recall that the U.S. stock market gained on average approx. 12% p.a. since 1926 with a return volatility of 20%. An investor who targets this 12% return holding a mean-variance efficient portfolio of long-only positions in stocks, bonds, and Rolex watches would invest 60.99% in stocks and 39.01% in Rolex, while holding no bonds at all. The resulting portfolio achieves a Sharpe ratio of 0.88 by reducing volatility to 11.41%, nearly half of the historic stock market volatility. Using only stocks and bonds for comparison, the Sharpe ratio of the efficient 76.87%/23.13% stock/bond portfolio amounts to only 0.70 because of higher return volatility of 14.30%. Again, these results indicate that luxury watches are beneficial for an investor, mostly by reducing risk even among well-diversified portfolios of common asset classes.

Jacobs et al. (2014) document that investors can benefit from using simple and cost-efficient time-invariant asset allocation policies and suggest weights of 60% for stocks, 25% for bonds, and 15% for commodities (here: gold), to obtain a suitable diversified world market portfolio.[13] If we follow their suggestion, the world portfolio generates an average return of 0.22% and an annualized Sharpe ratio of 0.96. It seems reasonable to analyze the impact of luxury watches on this combined world portfolio, because according to Deloitte (2023), typical investors in luxury goods and collectibles already hold combinations of these more traditional asset classes. We find that adding luxury watches to the world market portfolio has tremendous improvements for portfolio risk metrics. Minimizing the value at risk (VAR), one should hold 91.48% Rolex and only 8.52% in the world market portfolio, achieving an annualized 5%-VAR of only -2.23%.[14] Without luxury watches, the 5%-VAR would have been -27.17% and the conditional 5%-VAR -43.43%. According to Deloitte (2023), actual allocation to art and collectibles was on average 10.9%, based on a 2023 survey among wealth managers.[15] Based on these numbers, the 5%-VAR reduces to -24.44% for a combined investment of 5% Rolex and 95% world market portfolio, resp. -21.85% for a 10% Rolex weight. Similarly, the conditional 5%-VAR shrinks to -39.17%

[13]They do not consider real estate in their world market portfolio because investors are often already heavily exposed to real estate risk as outlined in Campbell (2006).

[14]Results are similar for other luxury watch indices.

[15]Private banks report an average of 8.6% allocation, and family offices report an average allocation of 13.4% to art and collectibles. In absolute figures, global UHNWI's (ultra-high net worth individuals) wealth associated with art and collectibles was $2.174 trillion in 2022.

(5% Rolex), resp. -35.15% (10% Rolex). Clearly, investments in luxury watches enhance the risk profile of already well-diversified portfolios of traditional asset classes.

Spanning tests

For our spanning tests, we follow the approach and notation of Huberman and Kandel (1987), Ferson and Keim (1993), and Kan and Zhou (2012). The main question addressed by these tests is whether the inclusion of additional test assets N improves an investor's investment opportunity in comparison with a set of already held benchmark assets K. Define R_{1t} as the K-vector of the returns on the K benchmark assets and R_{2t} as the N-vector of the returns on the N test assets. The regression specification including the test and benchmark assets under study is described by

$$R_{2t} = \alpha + \beta R_{1t} + \epsilon_t, \tag{5.3}$$

with R_{2t} representing either the return vector of a single luxury watch under study or the return matrix of a portfolio of all luxury watches combined. Because R_{1t} comprises the returns of our three benchmark assets stocks, bonds, and gold, the regression equation can be written as

$$R_{2t} = \alpha + \beta_{\text{stocks}} R_{\text{stocks}} + \beta_{\text{bonds}} R_{\text{bonds}} + \beta_{\text{gold}} R_{\text{gold}} + \epsilon_t. \tag{5.4}$$

As shown in Huberman and Kandel (1987), the necessary and sufficient conditions for spanning are given by the null hypothesis

$$H_0 := \alpha = 0_N, \quad \delta = 1_N - \beta 1_N = 0_N, \tag{5.5}$$

with 0_N and 1_N being N-vectors of zeros and ones, respectively. If the null hypothesis holds, we can find a portfolio of the K benchmark assets that has the same mean return but a lower variance than the test assets, hence the test assets are dominated. Our first test statistic LR is a likelihood ratio test of Eq. (5.5) that compares the likelihood functions under the null and the alternative. Besides LR, we follow Kan and Zhou (2012) and use the standard Wald test statistic W and the Lagrange multiplier test statistic LM. As shown in Berndt and Savin (1977) and Breusch (1979), these three tests impose $W \geq LR \geq LM$ in finite samples and could give conflicting results, with LM favoring acceptance and W favoring rejection. Next, the null hypothesis stated in Eq. (5.5) is a joint hypothesis. To better distinguish the effects of the N test assets on the global minimum-variance portfolio and the tangential portfolio of the K benchmark assets, we conduct the step-down test introduced in Kan and Zhou (2012). The first test statistic F_1 only tests $H_0 := \alpha = 0_N$, i.e., if the two tangential portfolios formed by K and $N + K$ differ. The second test statistic F_2 then testes $H_0 := \delta = 0_N$. If this null hypothesis is rejected, the two global minimum variance portfolios are statistically distinct. Finally, we use the generalized method of moments (GMM) estimation introduced by Hansen (1982) to account for the possibility that ϵ_t is conditional heteroscedastic. If so, our test statistics W, LM, and LR would not be viable. As they all have the same form under GMM estimation, we only report the according robust Wald test statistic W_{GMM} (see Kan and Zhou (2012)).

Table 5.5: Tests of mean-variance spanning.

This table presents two sets of mean-variance spanning tests on six luxury watch brand portfolios, using the value-weighted U.S. market return (stocks), the S&P U.S. Treasury Bond Index (bonds), and gold (LBMA p.m. fixing) as benchmark assets. The first three columns refer to the Wald (W), likelihood ratio (LR), and Lagrange multiplier (LM) test statistics for the null hypothesis $H_0 : \alpha = 0_N, \delta = 0_N$ (see the main text for details). The next two columns are the results for the step-down test, where F_1 tests for an improvement of return (i.e., $H_0 : \alpha = 0_N$) and F_2 for a possible risk reduction ($H_0 : \delta = 0_N$, conditional on $\alpha = 0_N$). The last column shows the Wald statistic for the generalized method of moments (W_{GMM}) accounting for conditional heteroskedasticity. All tests are performed on each luxury watch brand portfolio as well as jointly on all six luxury watch brand portfolios. Reported in brackets are p-values which are exact under the normality assumption on the residuals. The GMM Wald test has an asymptotic χ^2_{2N} distribution, where N is the number of test assets (here: $N = 3$), and related p-values are asymptotic ones. The results are presented for the entire sample period as well as for its two subperiods. The sample period begins 01/02/2017 for Rolex, Tudor, and Omega, 01/02/2019 for Audemars Piguet and Patek Philippe, and 01/02/2020 for Cartier and ends 09/30/2024.

	W	LR	LM	F_1	F_2	W_{GMM}
Panel A: Overall period, 01/02/2017 to 09/30/2024						
Audemars Piguet	224.60	167.53	128.26	22.45	185.70	202.86
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Cartier	512.28	277.38	166.65	0.01	506.04	803.14
	[0.00]	[0.00]	[0.00]	[0.93]	[0.00]	[0.00]
Omega	331.71	242.17	182.15	3.82	322.35	275.97
	[0.00]	[0.00]	[0.00]	[0.05]	[0.00]	[0.00]
Patek Philippe	200.45	153.41	120.00	14.75	174.89	155.43
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Rolex	438.17	296.77	210.20	24.20	387.23	317.51
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Tudor	411.81	283.92	203.93	0.04	408.68	419.72
	[0.00]	[0.00]	[0.00]	[0.85]	[0.00]	[0.00]
All watches	926.52	401.34	214.24	3.60	145.61	1489.25
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Panel B: First sub-period, 01/02/2017 to 03/07/2022						
Audemars Piguet	224.84	141.87	95.16	142.48	41.05	215.09
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Cartier	120.45	81.99	58.30	2.15	112.86	239.16
	[0.00]	[0.00]	[0.00]	[0.15]	[0.00]	[0.00]
Omega	114.56	95.49	80.43	12.36	96.39	66.40
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Patek Philippe	197.98	130.09	90.00	109.17	50.38	123.22
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Rolex	314.35	208.46	145.25	120.58	130.62	171.92
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Tudor	142.35	114.34	93.21	9.86	126.29	135.54
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
All watches	571.51	247.28	136.60	33.70	35.52	788.59
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Panel C: Second sub-period, 03/14/2022 to 09/30/2024						
Audemars Piguet	522.60	212.96	106.65	75.06	275.94	439.07
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Cartier	473.11	202.46	104.42	2.45	451.54	575.34
	[0.00]	[0.00]	[0.00]	[0.12]	[0.00]	[0.00]
Omega	365.56	176.33	98.06	6.34	334.67	370.54
	[0.00]	[0.00]	[0.00]	[0.01]	[0.00]	[0.00]
Patek Philippe	325.61	165.16	94.93	50.70	192.26	300.89
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Rolex	662.23	238.79	111.45	77.05	357.74	656.05
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Tudor	513.77	211.14	106.28	35.57	366.23	411.84
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
All watches	1371.05	361.62	157.70	22.00	130.79	1753.00
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]

Table 5.5 summarizes the results for the spanning tests over the full sample period from 01/02/2017 to 09/30/2024 (Panel A) and for two consecutive sub-periods (Panel B and C). Under both normality and non-normality assumptions, as well as overall and for both sub-periods separately, the null hypothesis of spanning can be rejected for all single luxury watch brand portfolios at a level of 1%. While this result is not surprising given that all luxury watch portfolios lie outside the efficient frontier formed by our three benchmark assets stocks, bonds, and gold (see Figure 5.3), it rises an important question for any investor: What is the source of improvement when adding luxury watches to already well-diversified portfolios? The fourth column of Table 5.5 shows that the test statistics F_1 cannot be rejected for Cartier, Tudor, and Omega at least at the 5% level. This implies that the diversification benefit of these brands is not attributable on the improvement of portfolio returns (i.e., an improvement of the tangency portfolio). On the other hand, Audemars Piguet, Patek Philippe, and Rolex seem to improve portfolio returns which is in line with our previous results that these brands are the ones with highest returns among all luxury watches (see Table 5.2). In addition, all p-values for the test statistic F_2 are throughout near zero, implying that risk-reduction is a strong source of diversification benefits of luxury watches. All luxury watch brand portfolios improve the global minimum-variance portfolio at a significance level of at least 1%. Given the descriptive statistics of luxury watch returns, this result is, however, not surprising. Compared with stocks, bonds, and gold, luxury watch returns have very low risk in form of standard deviation and display very low correlation with these asset classes. In summary, we find that an investor with an existing portfolio comprising stocks, bonds, and gold can expand her opportunity set by investing in luxury watches. While all luxury watches under study are beneficial in form of risk reduction, adding Audemars Piguet, Patek Philippe, and Rolex contemporaneously enhances portfolio returns.

Table 5.6: Spanning regressions.

This table reports estimated coefficients from the following time-series regression:

$$r_{i,t} = \alpha_i + \beta_{1,i} L4(\text{Stocks}_t) + \beta_{2,i} L4(\text{Bonds}_t) + \beta_{3,i} L4(\text{Gold}_t) + \epsilon_{i,t},$$

where $r_{i,t}$ is the weekly return on a given luxury watch brand portfolio i. Stocks denotes the value-weighted U.S. market return, bonds the return of the S&P U.S. Treasury Bond Index, and gold the return according to the LBMA p.m. fixing prices. $L4$ transforms a variable into a row vector consisting of four lags of that variable. Newey and West (1987) corrected standard errors are applied to compute robust t-statistics and */**/*** indicate that estimated coefficients are significant at a level of 10%/5%/1%. The sample period begins 01/02/2017 for Rolex, Tudor, and Omega, 01/02/2019 for Audemars Piguet and Patek Philippe, and 01/02/2020 for Cartier and ends 09/30/2024.

	(1)		(2)		(3)		(4)		(5)		(6)	
	Audemars Piguet		Cartier		Omega		Patek Philippe		Rolex		Tudor	
Variables	coef.	*t*-stat	coef.	*t*-stat	coef.	*t*-stat	coef.	*t*-stat	coef.	*t*-stat	coef.	*t*-stat
α	0.21**	2.17	0.00	−0.12	0.03	0.85	0.19*	1.76	0.10*	1.79	−0.01	−0.36
Stocks_t	−0.01	−0.76	0.00	−0.23	0.02*	1.74	0.00	−0.10	−0.01	−1.01	−0.01	−0.83
Bonds_t	−0.02	−0.25	−0.03	−0.86	0.06	1.14	−0.20**	−2.02	−0.03	−0.45	0.05	1.08
Gold_t	0.02	0.72	0.00	−0.07	−0.01	−0.50	0.05	1.25	0.03	1.45	0.00	−0.21
Stocks_{t-1}	0.00	0.27	0.01	1.29	0.02**	2.34	0.01	0.51	0.01	0.45	0.01	1.42
Bonds_{t-1}	−0.04	−0.46	0.05	1.14	0.03	0.60	−0.02	−0.25	0.00	−0.04	0.02	0.41
Gold_{t-1}	−0.01	−0.34	0.00	0.34	−0.02	−1.26	0.00	0.04	0.02	1.17	0.01	0.50
Stocks_{t-2}	0.01	0.75	0.01	0.97	0.01	1.02	−0.01	−0.57	0.02*	1.75	0.00	0.45
Bonds_{t-2}	−0.05	−0.74	0.06	1.52	0.01	0.32	−0.04	−0.35	0.01	0.17	0.03	0.54
Gold_{t-2}	−0.02	−0.59	−0.03*	−1.91	0.01	0.57	0.02	0.50	0.01	0.84	0.01	0.92
Stocks_{t-3}	0.01	0.54	0.01	0.82	0.03***	2.66	0.00	0.06	0.02	1.30	0.01	0.51
Bonds_{t-3}	−0.03	−0.40	0.12***	2.79	−0.01	−0.32	−0.03	−0.32	0.06	1.11	0.02	0.38
Gold_{t-3}	0.02	0.87	−0.02	−1.27	0.01	0.85	0.02	0.52	0.01	0.77	0.01	0.85
Stocks_{t-4}	0.01	0.73	0.01	1.06	0.00	−0.04	−0.01	−0.45	0.02*	1.82	−0.01	−1.12
Bonds_{t-4}	−0.05	−0.72	0.06	1.40	−0.01	−0.16	−0.01	−0.13	−0.01	−0.20	0.01	0.29
Gold_{t-4}	−0.02	−0.72	0.02	1.14	0.03*	1.93	0.01	0.37	0.02	1.18	0.01	0.73
Sum $t-1$ to $t-4$ stocks	0.04		0.04		0.06		-0.01		0.06		0.01	
Sum $t-1$ to $t-4$ bonds	-0.17		0.29		0.02		-0.10		0.05		0.08	
Sum $t-1$ to $t-4$ gold	-0.02		-0.03		0.03		0.04		0.06		0.05	
	$\chi^2(1)$	p-value	$\chi^2(1)$	p-value	$\chi^2(1)$	p-value	$\chi^2(1)$	p-value	$\chi^2(1)$	p-value	$\chi^2(1)$	p-value
χ^2 (1)[Sum $t-1$ to $t-4$ stocks=0]	0.46	0.50	3.62	0.06	7.15	0.01	0.01	0.92	2.72	0.10	0.47	0.49
χ^2 (1)[Sum $t-1$ to $t-4$ bonds=0]	0.55	0.46	11.58	0.00	0.05	0.82	0.11	0.74	0.09	0.76	0.50	0.48
χ^2 (1)[Sum $t-1$ to $t-4$ gold=0]	0.06	0.80	0.64	0.43	1.03	0.31	0.16	0.69	1.64	0.20	1.38	0.24
χ^2 (1)[α=0 and sum Betas=1]	17.51	0.00	60.23	0.00	78.85	0.00	10.81	0.00	15.68	0.00	60.08	0.00
N	299		247		404		299		404		404	
Adj. R^2	<0		0.04		0.02		<0		0.02		<0	
Annualized Alpha (in %)	10.87		-0.23		1.52		10.11		5.28		-0.60	

Finally, we present results of spanning regressions for robustness and to judge if any benchmark asset contributes to the explanation of average returns of luxury watches. To be specific, we regress weekly returns of luxury watch indices r_t on contemporaneous and up to four weeks lagged returns of stocks, bonds, and gold:

$$r_t = \alpha + \beta_1 L4(\text{Stocks}_t) + \beta_2 L4(\text{Bonds}_t) + \beta_3 L4(\text{Gold}_t) + \epsilon_t, \tag{5.6}$$

where $L4$ transforms a variable into a row vector consisting of four lags of that variable. As shown in Table 5.6, the intercepts of Cartier, Omega, and Tudor are not distinguishable from zero, which is unsurprising given that their average returns are close to zero. Again, these brands do not improve the returns of already well-diversified portfolios. On the other hand, estimated alphas are a large 5.28% p.a. for Rolex and exceed even 10% p.a. for Patek Philippe and Audemars Piguet. In addition, most coefficient estimates are statistically insignificant implying that returns of luxury watches are not well estimated by the returns of stocks, bonds, and gold. The sum of lagged coefficient estimates for stocks does not exceed 0.06 is close to zero for gold. The only exception is Cartier, where the sum of lagged coefficients for bonds adds up to 0.29, and Audemars Piguet, with a sum of -0.17, respectively. While there seems to be no clear economic channel between luxury watches and bonds, these results seem to reflect the fact that luxury watch returns have very low volatility in the same magnitude as bonds. As shown by the χ^2-tests, the null hypothesis that the sum of lagged coefficient estimates for stocks, bonds, and gold, is zero is not rejected in all cases but Cartier with respect to bonds. Looking at coefficient estimates for contemporaneous returns, all are close to zero and statistically insignificant with the exception of the bond coefficient estimate of -0.20 for Patek Philippe. Similar with our previous results, we reject the joint null hypothesis that alpha equals zero and the sum of all betas equals one for all six luxury watch brand portfolios. Overall, none of our benchmark returns shows covariation with luxury watch returns and there seems to be no clear return pattern up to a lag of four weeks.

Luxury watches and collectibles in general are a rather unusual investment class, so would a risk-averse individual really prefer the higher expected return of luxury watches over the lower expected return of traditional Treasury bills? This is more than an academic dispute because the practical implications of a difference in expected returns depend on whether risk is able to explain it. If not, public authority may design policies to eliminate these frictions to improve market efficiency. As for potential investors, if risk cannot explain differences in expected returns, they might avoid investments in luxury watches at all. To assess this risk-return relationship we apply the model-free test presented in Holcblat et al. (2024) to evaluate the null hypothesis if the investment in luxury watches strongly dominates a risk-free investment in Treasury bills in the sense of a second-order stochastic dominance. The basic idea behind this test is to check whether every possible risk-averse individual strictly prefers luxury watch returns over Treasury bill returns. We clearly reject this null hypothesis for all (daily) luxury watch brand portfolio returns as all resulting p-values are close to zero, so we conclude that any luxury watch investment appears to be associated with a higher exposure to risk.[16] In other words, levels of risk aversion exist s.t. luxury watches are preferred to Treasury bills. This implies that luxury watches are not a deviation from the risk-return tradeoff and rational, risk-avers investors would consider them as potential investments.

5.3.4 The impact of transaction costs

Transaction costs associated with buying and selling collectibles can be quite high. However, compared to oldtimers or wine investments, related costs for luxury watches are expected to be much lower because of their tiny physical dimensions and their much higher insensitivity to room temperature and humidity.[17] To develop after-cost estimates of average returns on luxury watches, we follow the approach proposed in Krasker (1979). Using our weekly data, we estimate the following regression using GLS:

$$\frac{p_{i,t+1} - p_{i,t}}{p_{i,t}} - rf_{t+1} = \sum_{i=1}^{N} \theta_i + \delta \frac{1}{p_{i,t}} + \frac{\epsilon_{i,t+1}}{p_{i,t}}, \tag{5.7}$$

where $p_{i,t}$ is the average price of luxury watch $i \in \{1, ..., N\}$ in week t, rf is the risk-free rate, and ϵ the residual. The resulting GLS estimate for the slope coefficient δ is 11.94, which implies that the costs for luxury watches are \$11.94 per week. To put this number in perspective, Krasker (1979) estimates the total costs of wine investments to be \$1.40 per bottle per year. Accounting for inflation, this equals \$7.25 in 2025-\$.[18] Unfortunately, Krasker (1979) does not show summary statistics for absolute prices, so the study remains silent about the costs relative to the absolute initial investment. Using 26,640 observed transactions of the most popular wines sold at auction at Christie's and Sotheby's between February 2007 and December 2013, the average price of 232 different wines is \$1,191.05

[16]This result is robust when analyzing sub-periods or weekly returns.
[17]Many luxury watches are quite robust and especially constructed for racing or scuba diving purposes.
[18]The sample in Krasker (1979) ends in December 1977. We take the CPI: All items less food and energy from FRED as a proxy for inflation.

(2025-\$).[19] Taken together, this implies transaction costs of 0.61% per bottle per year or 0.05% per month.

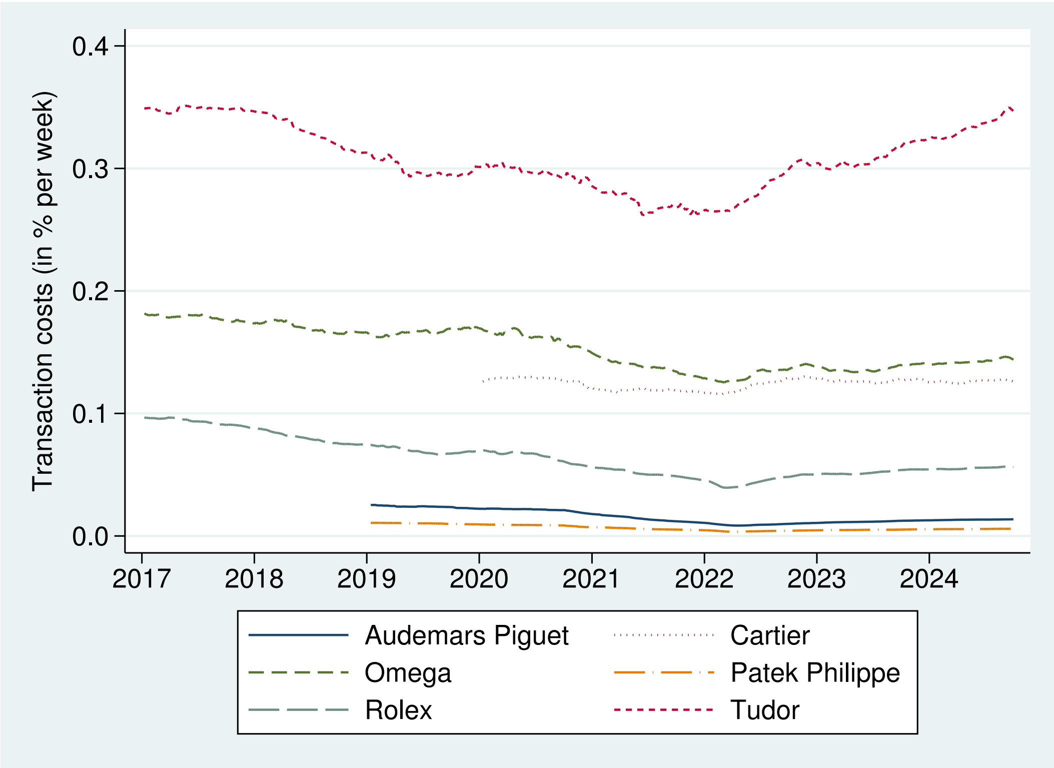

Fig. 5.4. This figure plots transaction costs for our six luxury watch portfolios using the cost estimation approach proposed in Krasker (1979). Transaction costs are presented relative to the average price of related luxury watches (in %) on a weekly basis. The sample period begins 01/02/2017 for Rolex, Tudor, and Omega, 01/02/2019 for Audemars Piguet and Patek Philippe, and 01/02/2020 for Cartier and ends 09/30/2024.

Our estimated cost of holding luxury watches of \$11.94 per week may seem too high to the reader, but given an average watch price of \$54,507.13, this amounts to a modest 1.14% per year (annualized).[20] Figure 5.4 visualizes estimated costs for our six luxury watch brands. Overall, average costs relative to watch prices (i.e., the investment in absolute dollars) range from 0.01% per week in case of Patek Philippe to 0.31% per week in case of Tudor.[21] We correct our average (gross) returns presented in Table 5.2 for these costs and report net returns in Table 5.7 for our full sample period and two sub-periods.

[19] See Bocart and Hafner (2015). We thank the authors for sharing their data.

[20] Average absolute prices for the luxury watches in our sample are \$19,736.52 for Rolex, \$85,274.27 for Audemars Piguet, \$9,608.36 for Cartier, \$7,869.99 for Omega, \$200,653.76 for Patek Philippe, and 3,899.90 for Tudor.

[21] Our findings are in line with evidence from the art market, where artwork with high common market value is traded more frequently, which implies lower transaction costs for higher priced items (see Nozari (2022)).

Table 5.7: Average weekly discrete returns of luxury watches net of transaction costs.

This table shows average weekly discrete returns (in %) for six luxury watch indices under study. Each luxury watch index comprises a portfolio of 30 watches of the related brand and their returns are weighted by market shares. Transaction costs (in %) are estimated using the method proposed in Krasker (1979). The Sharpe ratio is expressed annualized. T-statistics based on Newey and West (1987) corrected standard errors are provided in parenthesis. */**/*** indicates that coefficients are significantly different from zero at a level of 10%/5%/1%. The sample period begins 01/02/2017 for Rolex, Tudor, and Omega, 01/02/2019 for Audemars Piguet and Patek Philippe, and 01/02/2020 for Cartier and ends 09/30/2024.

	A. Piguet	Cartier	Omega	P. Philippe	Rolex	Tudor
Panel A: Overall period, 01/02/2017 to 09/30/2024						
Gross return	0.21	0.00	0.06	0.21	0.14**	0.00
	(1.62)	(0.06)	(1.59)	(1.51)	(1.96)	(0.09)
Transaction costs	0.02	0.12	0.15	0.01	0.06	0.31
Net return	0.20	−0.12***	−0.10***	0.20	0.07	−0.31***
	(1.51)	(−2.93)	(−2.60)	(1.46)	(1.04)	(−8.17)
Sharpe ratio	1.41	−2.58	−1.68	1.24	0.43	−4.69
Panel B: First sub-period, 01/02/2017 to 03/07/2022						
Gross return	0.64***	0.07	0.14***	0.66***	0.33***	0.10***
	(4.71)	(1.22)	(3.51)	(4.47)	(5.38)	(3.13)
Transaction costs	0.02	0.12	0.16	0.01	0.07	0.31
Net return	0.62***	−0.06	−0.03	0.65***	0.26***	−0.21***
	(4.53)	(−1.00)	(−0.64)	(4.41)	(4.14)	(−6.16)
Sharpe ratio	6.11	−0.83	−0.52	5.34	3.37	−2.93
Panel C: Second sub-period, 03/14/2022 to 09/30/2024						
Gross return	−0.32***	−0.05	−0.10*	−0.35***	−0.26***	−0.20***
	(−4.44)	(−0.98)	(−1.85)	(−4.59)	(−3.93)	(−3.93)
Transaction costs	0.01	0.13	0.14	0.00	0.05	0.31
Net return	−0.33***	−0.18***	−0.23***	−0.36***	−0.31***	−0.51***
	(−4.61)	(−3.38)	(−4.51)	(−4.66)	(−3.77)	(−10.54)
Sharpe ratio	−6.92	−4.66	−5.30	−5.45	−8.51	−11.23

For the full sample period, accounting for costs results in significant negative returns for Cartier, Omega, and Tudor. Unsurprisingly, the negative performance of all luxury watches since 03/14/2022 worsens if we consider costs. While this may be meaningful for a serious investor, some owners of these watches may still prefer them for some non-monetary dividend yield considering the aesthetic pleasure and the social status of the owner (Goetzmann (1993)).

Looking at the first sub-period until 03/07/2022, Audemars Piguet, Patek Philippe, and Rolex still earn large net returns. While Rolex generates a net return of 0.26% per week, it is 0.62% in case of Audemars Piguet and even 0.65% for Patek Philippe. To be rather conservative, if we further assume additional costs in form of fees for buying/selling watches amounting to a fixed 2.5%, investing in Rolex would still be profitable after the short period of less than 10 weeks, and in case of Audemars Piguet and Patek Philippe, these fees are even covered in less than one month.[22] In conclusion, despite the costs associated with luxury watch investments, which may be quite high, some luxury watches yield highly significant net returns between 01/02/2017 and 03/07/2022, in particular Rolex (0.26%), Patek Philippe (0.65%), and Audemars Piguet (0.62%).

5.4 Day-of-the-week effects

5.4.1 Overview

Our main finding so far is that luxury watches are beneficial for investors by enhancing portfolio returns (Audemars Piguet, Patek Philippe, and Rolex) and reducing risk (Cartier, Omega, and Tudor). To benefit from this diversification potential, an investor has to buy luxury watches, so the question arises when is the best time, in the sense of which day of the week, to buy them? Luxury watches can be traded all seven days of the week instead of just five days as for stocks. Similarly, cryptocurrencies are also continuously traded and Dorfleitner and Lung (2018) find that their returns are significantly lower an Sundays. In effect, investors can optimize their investment performance by timing their transaction accordingly. In this section, we address this key question and investigate if there exist seasonal patterns in daily returns of luxury watches.[23]

[22] While a fixed fee of 2.5% may seem to be rather arbitrary, our reference point is PayPal. Buyers may prefer paying through online services like PayPal, since one is eligible for buyers' protection program if paid this way. PayPal charges 1.9%-2.9% for this service.

[23] Seasonal return patterns are vastly documented in previous studies with regards to stocks. French (1980) find a day-of-the-week (DOW) effect, i.e., returns are smaller and even negative on Mondays. Ariel (1987) and Lakonishok and Smidt (1988) find that stock returns are significantly positive at the turn of the month, i.e., from the last trading day of a month until the third trading day of the following month.

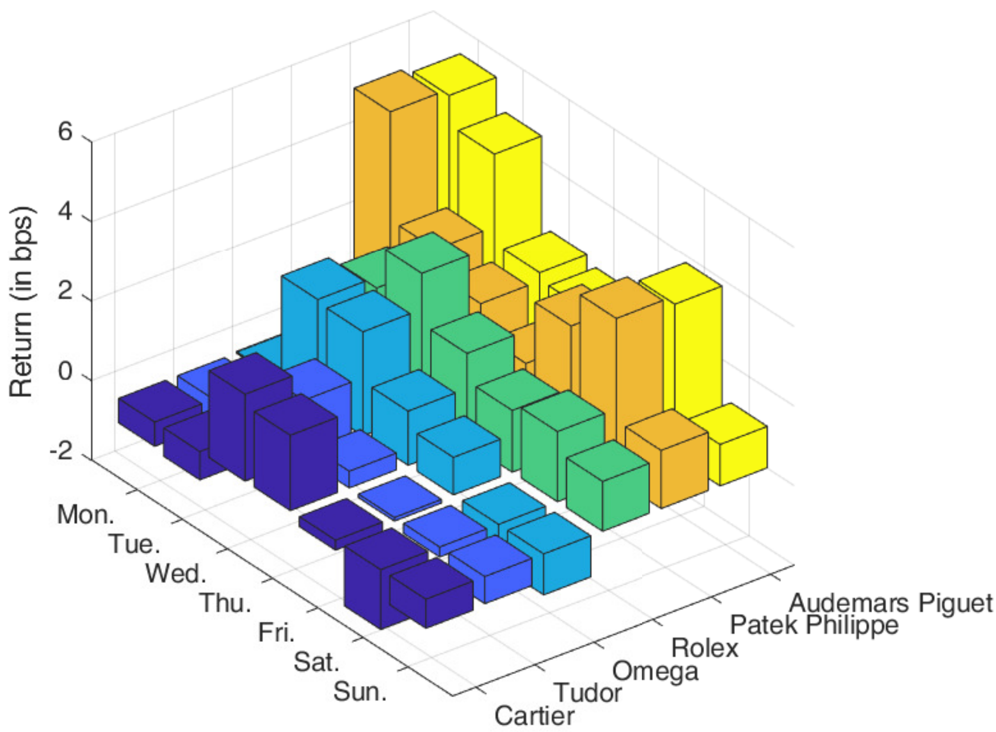

Fig. 5.5. Daily discrete returns (in basis points) of luxury watch brand indices until 09/30/2024. The sample period begins 01/02/2017 for Rolex, Tudor, and Omega, 01/02/2019 for Audemars Piguet and Patek Philippe, and 01/02/2020 for Cartier.

To begin with our analysis of potential day-of-the-week effects, Figure 5.5 visualizes average (discrete) returns throughout the week and we immediately notice that returns on weekends seem to be lower than on other days of the week. To test for differences in averages returns across the days of the week, we apply the following regression:

$$r_{j,t} = \sum_{i=1}^{7} \alpha_{i,j} D_{i,t} + \epsilon_{j,t}, \tag{5.8}$$

where $r_{j,t}$ is the simple return of luxury watch brand portfolio j on day t and $D_{i,t}$ are dummy variables for each day of the week. Thus, $D_{1,t}$ equals 1, if the return is observed on a Monday and 0 otherwise, $D_{2,t}$ equals 1 on Tuesdays and 0 otherwise, and so on. Because we omit an intercept in Eq. (5.8), we include dummy variables for $i = 1, ..., 7$, i.e., from Mondays ($i = 1$) to Sundays ($i = 7$). Estimated coefficients $\hat{\alpha}_{i,j}$ measure the average return of luxury watch brand portfolio j on day i, which are summarized in Table 5.8. To assess the statistical significance of coefficient estimates, we use White (1980) corrected standard errors. In addition to the regressions for each brand, we report coefficient estimates for the pooled regression using data on all brands in the first column of Table 5.8. In the pooled regression, we cluster standard errors by day and brand.

On Saturdays and Sundays, returns for the six brand indices are on average an insignificant and small 1.12 bps, resp. 0.18 bps. Highest returns are observed on Wednesdays with 2.42 bps, which is mostly driven by the economically large and statistically highly significant Wednesday return of 3.51 bps for Rolex. In general, returns tend to be higher around the mid of a week with average returns of 2.12 bps on Tuesdays and 1.55 bps on Thursdays (both driven by high returns of Audemars Piguet and Rolex). At the brand level, Cartier watches yield negative returns on most days but generate a remarkable 2.18 bps on Wednesdays. Tudor returns are negative from Fridays to Mondays but generate a relatively high 1.46 bps on Wednesdays. Overall, we reject the null hypothesis of equal returns across the days of the week for all luxury watch brands except Cartier and Tudor. Our results are similar with the findings in Dorfleitner and Lung (2018) for cryptocurrencies, where Sundays returns are significantly lower than those on other days. However, Aharon and Qadan (2019) documents that Mondays are associated with higher returns for Bitcoin, whereas we find higher returns around Wednesdays.[24]

5.4.2 Conditional variance framework

Section 5.3 demonstrates that investors should consider luxury watches as an alternative investment with attractive risk-return properties like low correlation with stocks, bonds, and gold. This raises the question when to buy luxury watches? This question is more crucial for luxury watches than stocks, because they can be traded on two more days, i.e., Saturdays and Sundays. The focus of the previous section has been on patterns in average returns of six luxury watch brand portfolios throughout the days of the week and we find relatively low returns on Sundays. However, an investor does not only consider returns in an investment decision-making process, but focuses on the volatility of returns, too. For that reason, it is important to know if a high (low) return is associated with a correspondingly high (low) volatility for a given day. To provide an example of the dynamic behavior of volatility, we take the (annualized) rolling 30-day volatility of returns. In case of Rolex, the average volatility of 2.29% varies between 0.82% (10/24/2023)

[24]Comparing these findings with seasonal patterns found in equity returns (French (1980)) is not meaningful because related studies do not account for the trading opportunity on weekends.

Table 5.8: Day-of-the-week effects.

The first column of this table reports estimates from a regression of luxury watch brand portfolio returns (in basis points) on dummy variables for each day of the week. We estimate this regression using pooled data and cluster standard errors by day and brand (related t-statistics are presented in parenthesis). The remaining columns show average returns (in basis points) of luxury watch brand portfolios by day of the week and related robust t-statistics according to White (1982) in parenthesis. */**/*** indicates statistical significance at a level of 10%/5%/1%. Presented in brackets is the percentage rate of positive days, i.e., $r_{j,t} > 0$ for luxury watch brand j conditional on the day of the week. Significance levels are based on a sign test with the null hypothesis that the probability of a positive return is 50%. F-statistic refer to an overall F-test with the null hypothesis that average returns are equal across days of the week. The sample period begins 01/02/2017 for Rolex, Tudor, and Omega, 01/02/2019 for Audemars Piguet and Patek Philippe, and 01/02/2020 for Cartier and ends 09/30/2024.

	All	Rolex	P. Philippe	A. Piguet	Cartier	Omega	Tudor
Monday	1.42	0.72	5.52***	5.36***	−0.61	0.04	−1.20
	(1.22)	(1.06)	(4.54)	(4.25)	(−0.59)	(0.03)	(−1.33)
	[44.01***]	[47.65]	[44.67*]	[51.00]	[44.76]	[43.21]***	[35.06***]
Tuesday	2.12***	2.40**	2.81**	4.63***	−0.71	2.68**	0.64
	(3.01)	(3.30)	(2.29)	(4.57)	(−0.58)	(2.42)	(0.61)
	[45.55***]	[51.24]	[41.47***]	[43.81**]	[46.56]	[50.25]	[38.86***]
Wednesday	2.42***	3.51***	2.18*	2.39**	2.18*	2.61	1.46
	(4.76)	(4.63)	(1.68)	(2.31)	(1.67)	(1.62)	(1.58)
	[46.77***]	[59.16***]	[37.67***]	[43.67**]	[53.04]	[44.80**]	[41.58***]
Thursday	1.55***	2.24**	1.42	2.21**	1.89*	1.36	0.42
	(3.91)	(3.14)	(1.37)	(2.02)	(1.86)	(1.21)	(0.47)
	[47.33**]	[57.18***]	[36.00***]	[44.67**]	[52.02]	[50.50]	[41.83***]
Friday	1.09**	1.55**	3.10***	1.39	−0.25	0.94	−0.09
	(2.13)	(2.03)	(2.77)	(1.52)	(−0.25)	(0.84)	(−0.11)
	[44.71***]	[51.49]	[41.33***]	[42.00***]	[45.56]	[46.29]	[40.35***]
Saturday	1.12	1.76**	4.03**	3.82***	−1.50	−0.71	−0.25
	(1.16)	(2.49)	(2.11)	(3.12)	(−1.12)	(−0.60)	(−0.30)
	[45.29***]	[51.24]	[42.67**]	[43.67**]	[47.58]	[46.53]	[39.85***]
Sunday	0.18	1.25*	1.46*	1.04	−0.73	−1.03	−0.68
	(0.37)	(1.90)	(1.67)	(1.50)	(−0.68)	(−0.93)	(−0.70)
	[40.73***]	[51.98]	[38.00***]	[39.67***]	[43.55**]	[39.11***]	[32.18***]
N	14,419	2829	2099	2099	1734	2829	2829
F-statistic	2.45	8.51	6.39	10.13	1.29	1.74	0.81
p-value	0.02	0.00	0.00	0.00	0.25	0.10	0.58

and 5.51% (02/09/2019). In other words, the relatively high volatility of 5.51% is more than three standard deviations above its mean, given that the volatility of rolling 30-day volatility is 1.04% for Rolex returns.

Using a conditional variance framework as a robustness test, we reexamine our findings for day-of-the-week effects and explicitly account for the dynamic behavior of watch return volatility by applying an exponential ARMA-EGARCH(-M) model (Engle (1982), Nelson (1991)) in this section.[25] A major advantage of this specific type of GARCH model is its ability to allow for asymmetries in shocks to the conditional variance typically found in financial time-series: Negative return shocks tend to increase volatility more than positive shocks, which is known as the leverage effect (Black (1976)). To be specific, the model to analyze the day-of-the-week effect in this section is specified by the conditional mean equation

$$r_t = c + \sum_{i=1}^{6} \delta_i D_{i,t} + \sum_{j=1}^{s} \phi_j r_{t-j} + \sum_{j=1}^{m} \varphi_j \epsilon_{t-j} + \eta \sigma_t + \epsilon_t, \tag{5.9}$$

and the conditional variance equation

$$\ln\left(\sigma_t^2\right) = \omega + \sum_{i=1}^{6} \delta_i D_{i,t} + \sum_{j=1}^{q} \left(\alpha_j z_{t-j} + \gamma_j \left(|z_{t-j}| - E[|z_{t-j}|]\right)\right) + \sum_{j=1}^{p} \beta_j \ln\left(\sigma_{t-j}^2\right), \tag{5.10}$$

where r_t denotes the log-return on day t and σ_t^2 its variance.

In addition to the intercepts c and ω, the model contains dummy variables $D_{i,t}$ for each day of the week except Monday, to avoid the dummy variable trap. To be specific, D_1 equals the value one on Tuesdays and zero otherwise, D_2 equals one on Wednesdays and zero otherwise, and so on. $\epsilon_t = z_t \sigma_t$ represents the error process, where z_t is a random variable with mean zero and variance of one. The specific order of the ARMA(s,m) and GARCH(p,q) terms as well as the consideration of including the volatility of returns in the mean equation is based on minimizing the Akaike information criterion (AIC), Bayesian information criterion (BIC), and the Shibata information criterion (SIC). To find the optimal model specifications according to these criteria, a maximum order of GARCH(2,2) is predetermined and we consider the normal distribution (norm), the

[25]This model was previously used to investigate seasonal patterns for other asset classes, e.g., stocks (Abalala and Sollis (2015)), real estate (Bampinas et al. (2016)), or cryptocurrencies (Dorfleitner and Lung (2018)). The M refers to the potential consideration of GARCH effects within the autoregressive moving-average process.

Student distribution (std), their skew variants (snorm, resp., sstd), and the normal inverse gaussian distribution (nig), as valid choices for the distribution of z_t. Within the conditional variance equation, α_j displays the sign (leverage) effect and γ_j the size effect, i.e., the impact of larger deviations of innovations from their expected value. Within the conditional mean equation, η captures the impact of the standard deviation and for that reason illustrated the consideration of the GARCH process within the mean process.

Table 5.9 shows the results for day-of-the week effects using our conditional variance framework. Panel A (B) show estimated coefficients for the conditional mean (variance) equation for the six luxury watch brand portfolios under study. Insights for the suitability of the model are provided in the lower part of Table 5.9. Panel C reports several information criteria and diagnostic tests. First, the sign bias test (*SignBias*, see Engle and Ng (1993)) tests the presence of leverage effects in the standardized residuals to capture possible misspecifications of the GARCH model. We report test statistics for the null hypothesis that both negative and positive reactions to lagged shocks not already captured by the model jointly equal zero. Next, we report test statistics for the (weighted) ARCH-LM test as proposed by Fisher and Gallagher (2012). This is essentially a weighted version of the portmanteau test proposed by Li and Mak (1994) for testing the null hypothesis of an adequately fitted ARCH process. We use standardized squared residuals up to the seventh lag corresponding to a whole week. Panel D reports the test statistic $\widetilde{Q}(7)$ for the (weighted) Ljung-Box test as proposed by Ljung and Box (1978) and Fisher and Gallagher (2012) to test the null hypothesis of no serial correlation remaining for standardized residuals up to the seventh lag. Finally, the remaining rows in Panel D of Table 5.9 report skewness and excess kurtosis of standardized residuals.

To begin with the estimated coefficients for the conditional mean equation in Panel A, we observe that returns on Sundays are lower than on other days of the week. Because dummy variables for Sundays are omitted, this day provides the baseline for comparison. Only nine out of 36 dummy variable estimates $D_{i,t}$ are negative, but neither statistically significant at the 5% level, nor economically important. The negative Sunday effect is most pronounced for Omega, followed by Patek Philippe, and Tudor. In line with our previous results, we detect that returns for Rolex watches are significantly higher on Wednesdays, whereas Cartier returns show no pattern throughout the days of the week. Looking at Panel B of Table 5.9, we find conflicting results for day-of-the-week effects with regards to the volatility of returns. Coefficient estimates for Patek Philippe and Audemars Piguet are uniformly positive and statistically highly significant, thus indicating that according return volatility is much lower on Sundays. However, we observe lower volatility for Tudor

Table 5.9: Estimates of an ARMA-EGARCH(-M) model for day-of-the-week effects.

This table shows coefficient estimates of the day-of-the-week analysis for six luxury watch brand portfolios using a conditional variance framework. */**/*** indicates statistical significance at a level of 10%/5%/1% based on White (1982) corrected standard errors. The sample period begins 01/02/2017 for Rolex, Tudor, and Omega, 01/02/2019 for Audemars Piguet and Patek Philippe, and 01/02/2020 for Cartier and ends 09/30/2024.

	Rolex	P. Philippe	A. Piguet	Cartier	Omega	Tudor
Panel A: Mean equation						
c	1.5763	20.8981***	−0.6064***	−2.0670	−1.5451***	−0.5085***
δ_1 (Monday)	−1.2809*	0.8720*	0.4906	−1.5017	2.3212***	−0.1739
δ_2 (Tuesday)	0.1641	2.2172***	0.7219*	−1.6251	4.9624***	2.1787*
δ_3 (Wednesday)	1.6224***	0.1601	0.8160	1.1117	0.7111*	2.0554***
δ_4 (Thursday)	1.1643*	−0.6158	0.7180*	0.3630	2.9403***	2.2577**
δ_5 (Friday)	−0.0921	1.4082***	0.6292	−0.9944	2.7179**	0.5086***
δ_6 (Saturday)	−0.4517	10.0185***	0.6292	−0.8398	1.6976***	1.0536
ϕ_1	0.9123***	0.9910***	0.9941***	−0.1517***	1.2119***	
ϕ_2	0.0706***			−0.1357***	−0.9767***	
φ_1	−0.9012***	−0.9467***	−0.9888***	0.1054***	−1.2313***	
φ_2				0.0788***	0.9805***	
η	−0.0840	−0.2701***		0.1391		
Panel B: Variance equation						
ω	−0.2338	−1.5666***	−1.0140***	−0.6579**	−0.6400***	0.1191***
δ_1 (Monday)	0.1924	1.2209***	1.4767***	0.0955	0.2303	−0.2404
δ_2 (Tuesday)	0.4870*	2.0395***	1.8023***	0.2678	0.4157	0.4484
δ_3 (Wednesday)	0.4287*	1.1431***	1.1209***	0.4495	0.5782**	−0.6915**
δ_4 (Thursday)	0.0876	1.1791***	0.8532***	−0.1125	0.2182	−0.087
δ_5 (Friday)	0.1311	1.6511***	0.9651***	−0.0844	0.2589	−0.3260
δ_6 (Saturday)	0.0178	2.5525***	0.8669***	−0.0471	0.2767	−0.0176
α_1	−0.0016	0.0128	0.1019***	0.0452**	−0.0145	0.0879***
α_2	0.0130				0.0081	−0.0693***
β_1	0.0345***	0.9857***	0.9999***	0.4595***	0.3426***	0.9988***
β_2	0.9624***			0.4949***	0.6259***	
γ_1	0.0984***	0.0940***	−0.0350***	0.1778***	0.2714***	0.0227
γ_2	0.0053				−0.0282	−0.0098
skew	1.0831***	1.3596***	1.2316***	0.9686***	1.0412***	1.0939***
shape	3.4657***		2.0100***	3.5828***		
Panel C: Fit diagnostic tests						
Log-likelihood	15,321.99	10,454.14	11,589.94	8732.60	13,457.01	14,100.24
AIC	-10.81	-9.94	-11.02	-10.04	-9.50	-9.95
BIC	-10.76	-9.88	-10.97	-9.96	-9.44	-9.91
SIC	-10.81	-9.94	-11.02	-10.04	-9.50	-9.95
SignBias	5.68	0.49	3.31	1.51	3.74	0.26
ARCH-LM	1.85	0.12	1.20	1.53	0.79	1.99
Distribution (z_t)	sstd	snorm	sstd	sstd	snorm	snorm
Panel D: Standardized residuals $\frac{\epsilon_t}{\sqrt{h_t}}$						
$\widetilde{Q}(7)$	1.99	1.81	5.84	1.17	0.81	7.14
skewness	0.38	4.22	10.84	-0.90	0.18	1.22
exc. kurt.	5.54	57.44	222.92	10.70	7.00	16.90

on Wednesdays. While there is no evidence for a negative Sunday effect for Cartier, Omega, and Tudor, all coefficient estimates for Rolex are positive and with regards to Tuesdays and Wednesdays weakly significant. Overall, we find that returns for luxury watch brand portfolios are significantly lower than those on other days. The same effect effect is observed for their conditional variance, especially for the brands Patek Philippe and Audemars Piguet, but are less stable than those regarding average returns.

Estimated coefficients $\hat{\alpha}_j$ capture the sign (i.e., leverage) effect and are significantly positive for Audemars Piguet, Cartier, and Tudor. This implies a negative leverage effect, thus positive shocks which are beneficial for investors have a greater impact on the conditional variance than negative shocks. The (sum of) size estimates $\hat{\gamma}_j$ is positive for all luxury watch brand portfolio except Audemars Piguet, however, statistically insignificant for Tudor. These results suggest that greater deviations of innovations from their expected value have a greater impact on the conditional variance than smaller ones. Innovations are chosen from skewed distributions in all cases and the best model fit is achieved by choosing the skewed normal distribution for Patek Philippe, Omega, and Tudor, and the skewed Student distribution for Rolex, Audemars Piguet, and Cartier. Unsurprisingly, estimated skewness and shape coefficients are highly significant and highlight the highly non-normal distribution of luxury watch brand portfolio returns. This is further emphasized by the skewness and excess kurtosis of standardized residuals shown in Panel D of Table 5.9 which are different from zero in all cases. Finally, (the sum of) $\hat{\beta}_j$ estimates is highly significant and close to unity. This indicates a high level of persistence for the conditional variance. The ARCH-LM tests confirm the conclusion of no remaining ARCH effects, while results for the sign bias test show no remaining asymmetries caused by positive or negative shocks, that are not already accounted for by the model specification. The (weighted) Ljung-Box test concludes that there is no remaining serial correlation for standardized residuals up to the seventh lag.

Overall, our test statistics imply that the ARMA-EGARCH(-M) model fits the data of all six luxury watch brand portfolios well. We document that returns on Sundays are significantly lower than on other days. Similarly, a negative Sunday effect is also detected for the conditional variance of our luxury watch portfolios under study, however, it is less pronounced compared to the conditional mean.

5.4.3 Possible explanations

Sales volume

What explains lower returns and volatility of most luxury watch brand portfolios on Sundays?[26] A vast body of literature suggests that returns, volatilities, and trading volume covary among asset classes because of a common exposure to the rate of information flow into the market.[27] In case of cryptocurrencies, which are also traded on weekends, Koutmos (2018), Balcilar et al. (2017), and Dorfleitner and Lung (2018) find the low trading volume on Sundays to cause the negative Sunday effect of low returns. However, in contrast to stocks, cryptocurrencies, or other exchange traded financial products, collectibles are mostly traded on auctions, online platforms, or in private, so trading volume is hard to determine. As of 12/03/2024, our data provider WatchCharts Analytics reports that within past 12 months, 40,487 Rolex watches they tracked price information for were listed and actually sold on the online marketplace eBay. That makes eBay the platform with the highest sales volume for Rolex watches in our sample, followed at a large distance by Carousell (5,135 sold watches), Reddit (3,012), and RolexForums (2,674). For that reason, a first place to go for sales data is ChronoIndex, a website tracking daily global sales volume (in U.S.-$) from eBay for all popular Rolex models. Unfortunately, they disclose aggregated sales volumes only fragmented and for a maximum period of 90 days. After discharging missing data, we observe 70 daily changes (in %) for Rolex sales on eBay between 08/06/2024 and 11/27/2024.[28]

We observe relatively constant sales figures between \$359,409 and \$398,514 per day in line with statements from ChronoIndex that eBay generates around \$15 million in Rolex sales each month. The average sales volume is 1.96% higher on Sundays than on Saturdays whereas average sales volume increases by even 6.73% from Sundays to Mondays. Most striking is the remarkable increase in sales volume of 15.87% from Fridays to Saturdays, which is economically important although none of sales volume changes are significant on

[26]This section only refers to Rolex because of their high importance among luxury watches and data availability. Our analysis here is very limited because e.g., data on sales volume for Rolex watches on secondary marketplaces is only available for a very short time-period (90 days) and, to the best of our knowledge, generally not available for any other brand in our sample.

[27]See e.g., Karpoff (1987), Andersen (1996), Bessembinder et al. (1996), and Chen (2012) for stocks, Balduzzi et al. (2001) for bonds, Tsai (2014) for real estate, and Chiarella et al. (2016) for commodities. The hypothesis is often referred to as the "mixture of distributions hypothesis" first proposed by Tauchen and Pitts (1983).

[28]ChronoIndex states that they cover a maximum period of past 90 days but our data retrieval shows that most recent, non-missing 90 sales figures are returned, which explains the sample period longer than 90 days.

common statistical levels. This large increase in sales is, however, not accompanied by a significant return for Rolex watches on Saturdays, as shown in Table 5.9.

In summary, there seems to be no evidence that lower returns on Sundays for luxury watch brand portfolios are the outcome of lower trading, resp. sales, volume. In fact, the highest increase in Rolex sales actually occurs during the weekend. Following the argumentation in Dorfleitner and Lung (2018), the non-existing connection between trading volume (i.e., sales) and luxury watch returns on Sundays could be the outcome of a negative premium for being able to trade on weekends at all, which is a special advantage to investors compared with stocks or bonds. We investigate how the behavior of (potential) investors interacts via attention induced demand with contemporaneous portfolio returns in the next section.

Retail investor attention

Another important economic channel to be considered for Sundays' returns is investor attention.Barber and Odean (2008) claim that investors are net buyers of attention-grabbing stocks and that any increase in the attention paid to the market will temporarily pressure prices upwards.[29] Using search query data provided by Google Trends as a direct proxy for investor attention, Da et al. (2011), Joseph et al. (2011), Swamy and Dharani (2019), and Nguyen et al. (2020) provide supporting evidence for the proposed positive relation between stock returns and the Google Search Volume Index. Based on these insights, we use the daily change (in %) of the worldwide Google web search volume for the term "Rolex" as a proxy for investor attention for Rolex luxury watches. We collect these Google Trends data for our full sample period from 01/02/2017 to 09/30/2024.[30] Relative changes for the search term "Rolex" are illustrated in Figure 5.6 separately for each day of the week. The whiskers correspond to a 99% confidence interval according to White (1982) robust standard errors.

We find that Google web searches for the term "Rolex" remain quite constant from Tuesdays to Thursdays. Because we measure relative changes compared with the level of searches from the previous day, our findings imply that the term "Rolex" is a highly relevant query during the weekend. We observe that the search volume for term "Rolex" is 13.33% higher on Saturdays than on Fridays, and the higher level remains predominant

[29]See Ayala et al. (2024) for an extensive summary of the literature on the use of Google Trends data in financial research.

[30]While most studies on investor attention using Google Trends data analyze a monthly or weekly frequency, our approach is in line with the few existing studies applying a daily frequency, see e.g., Kristoufek (2015), Tang and Zhu (2017), or Smales (2021).

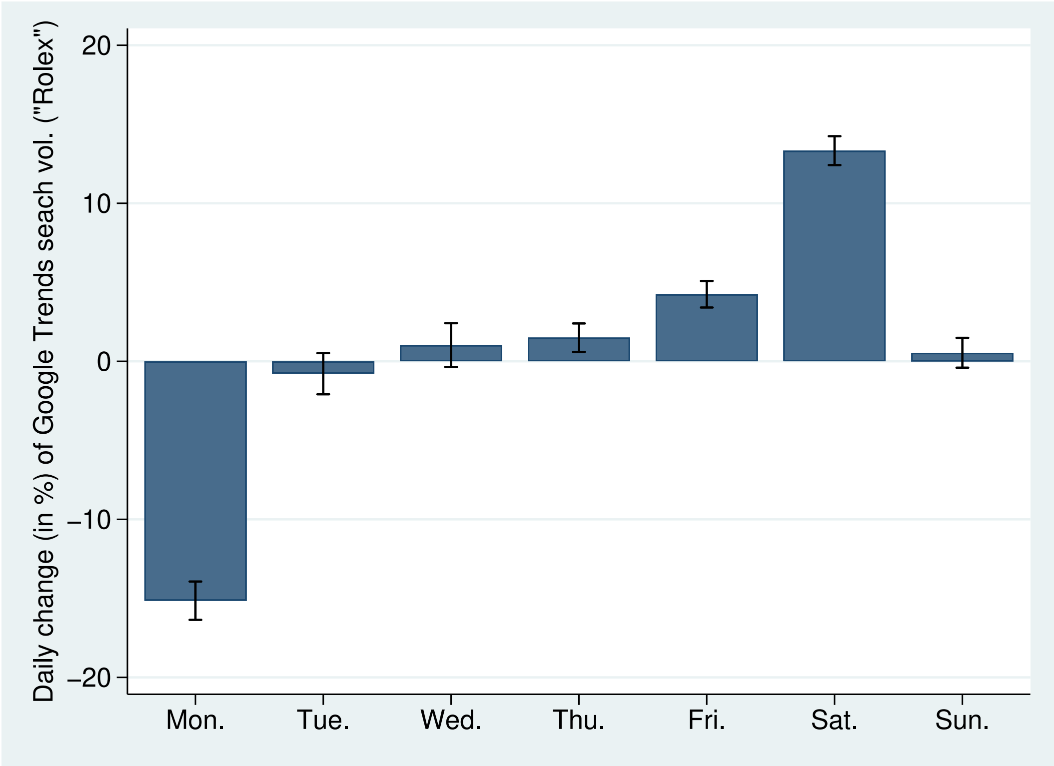

Fig. 5.6. This figure illustrates the daily change (in %) of the worldwide web search volume according to Google Trends for the term "Rolex" separately for each day of the week within our sample period from 01/02/2017 to 09/30/2024. The whiskers correspond to a 99% confidence interval according to White (1982) robust standard errors.

on Sundays.[31] What follows is a huge drop in Google searches from Sundays to Mondays of -15.15%. Da et al. (2011) argues that changes in investor attention measured by Google Trends data captures the behavior of individuals or retail investors instead of sophisticated traders or institutions. This may explain the huge search increase at the weekend, because the week is the source of most temporal organization, especially among full-time working individuals who like to take their time before investing large sums of money into hard to value collectibles like luxury watches. Based on the observed pattern in Google Trends searches, we would expect higher attention induced returns for the Rolex watch portfolio on Saturdays and Sundays, which is in contrast to our main findings. Although insignificant, the coefficient estimate for Saturday is even negative (-0.4517, see Table 5.9) when (relative) attention is at its peak.

The non-existence of the quite undisputed relation between attention and contempo-

[31]The search volume from Saturdays to Sundays only changes by +0.54%.

raneous returns found in many asset classes is very surprising, moreover because the possibility of trading luxury watches on weekends should be affected by findings of another huge body of literature on (individual) investor psychology: Asset returns are related to investors' mood which tends to increase from Monday through Friday.[32] Put simply, people tend to use their mood as the basis for forming evaluations and as mood is at its peak on Fridays throughout the week, overoptimistic projections drive prices (further) away from their fundamental value, especially for hard or highly subjective to value stocks. In result, speculative stocks earn high returns on Fridays and low returns on Mondays (Birru (2018)). Besides the stock market, investing in luxury collectibles requires the same decision-making process, and we detect that most individuals gather information about "Rolex" wristwatches on Saturdays and Sundays, when mood is near its peak. Again, the economic channel of investor mood implies higher returns of Rolex watch portfolio returns at the weekend, which is in contrast to our findings. This is even more surprising, given that wristwatches are naturally hard or highly subjective to value (which is a necessary assumption in the study of Birru (2018)), as their valuation depends on characteristics like condition, provenance, materials used, parts replaced, or the specific movement. Overall, our data rejects the hypothesis that attention induced demand and the behavior of individual investors causes lower returns on Sundays.

Seller behavior, price discovery, and the macroeconomy

In contrast with the continuous interaction of buyers and sellers on regulated markets like stock exchanges, luxury watches are offered by sellers in form of listings, so we generally observe ask-prices. This implies that observable market prices solely reflect expectations from sellers and potential buyers are not directly able to contribute to this price discovery. Our analysis on Google Trends search volume for the term "Rolex" discovered that most buyers may be retail, individual investors, but what do we know about the sellers? Looking at eBay (U.S.) as of 12/03/2024, there are 17,448 listings for Rolex watches with information about the seller warranty available. More than 80% of these offers (14,003) provide a warranty, mostly in form of a 14-day right to return. This return policy is generally granted by commercial sellers and not by individuals, so we conclude that the majority of luxury watch sellers on eBay are professional dealers. To

[32] See e.g., Rossi and Rossi (1977), Golder and Macy (2011), or Larsen and Kasimatis (1990) on day-of-the-week effects in mood. Birru (2018) provides evidence that Monday (Friday) alone accounts for over 100% of monthly returns for selected long-short portfolio strategies examined for which the short (long) leg is the speculative leg. Among these prominent strategies are the size-effect (Banz (1981)), illiquidity (Amihud (2002)), lottery-like characteristics as e.g., the Max-effect (Bali et al. (2011)), and profitability (Novy-Marx (2013) and Ball et al. (2015)).

add further credibility to this finding, we next analyze the dataset "Trending eBay Watch Listings 24" published by Kaggle (see Kanchana (2024)). The dataset contains a total of 2,000 watch listings from eBay's marketplace, reflecting the platforms' latest trends and preferences within the global watch community as of 2024. We use the timestamp of the last update to a listing to understand the behavior of luxury watch sellers, especially when at what time they decide to publish or update a listing. To begin with, we observe 1,850 listings with complete information about the timestamp and the offered price. eBay decided to display website user luxury watches as being trendy in the year 2024 that were initially published (or at least last updated) between 09/19/2023 and 04/01/2024. The average offered price for a luxury watch is \$2,876.20. The top two sellers "Watch County" and "Sigmatime" within this category amount for a total of 736 out of the 1,850 offers which further affirms our supposition that most offers stem from commercial sellers. Indeed, both are marked as top sellers and amount for a combined sales volume of more than 45,000 watches since 2011. If we look at the calendar day a listing was updated or published, we observe an interesting pattern: Most listings were updated or published on Thursday (464) and Tuesday (420), but only 74 out of 1,850 on Sundays. A Pearson's χ^2 test strongly rejects the null hypothesis that this observed difference between days of the week arose by chance. Interestingly, activity on Saturdays (313) is relatively high given that a uniformly distribution would result in $1,850/7 = 264.29$ edits per day.

To understand the timing behavior of luxury watch sellers on eBay in greater detail, we now take a look at the specific time of the day an offer is published or updated. Identification of related time clusters using k-means first requires to specify the number of clusters in advance. To mitigate concerns of choosing a number best suiting our findings, we use the gap statistic proposed by Tibshirani et al. (2001) to automatically determine the appropriate number of clusters. The algorithm suggests three clusters with average times within each cluster of 09:03 a.m., 05:03 p.m., and 01:07 a.m.[33] Most listings are assigned to the first (919) and second (544) cluster which basically match the begin and end of a typical business day.[34] Again, the time clusters further support our hypothesis that most sellers of luxury watches in our sample are professional dealers. Given the fact that Sunday is typically a day of rest in most western countries, it is no surprise that nearly all updates on Saturdays' offers (292 of 313) occur during the afternoon around

[33] All times are according to the Pacific Time Zone where most of the sellers are located. The ratio of between and total sum of squares is 91.7%. We use the algorithm proposed in Hartigan and Wong (1979) for k-means clustering.

[34] The remaining 387 listings that are edited overnight may be edited in advance and set to be published at a predefined time.

05:03 p.m. What follows are only 74 edits on Sundays with no clear pattern in time, and, after the resting day, 303 edits on Mondays, where 214 of them already occur within the morning cluster around 09:03 a.m.

To better understand the nature of our Rolex portfolio returns and for robustness of our findings in this section, we finally examine a large panel of macroeconomic and financial variables as candidate proxies. We use 126 economic series from Federal Reserve Economic Data (FRED) provided by McCracken and Ng (2016). These variables represent broad categories of the macroeconomy: (i) output and income, (ii) labor market, (iii) consumption, (iv) orders and inventories, (v) money and credit, (vi) interest and exchange rates, (vii) inflation, and (viii) stock market indicators. Because macroeconomic data is available on a monthly basis, we convert our Rolex portfolio returns for this analysis accordingly. To begin with, we regress $T = 98$ (Jan. 2017 - Sep. 2024) one-month ahead Rolex portfolio excess returns r^e_{t+1} on our panel of 126 macroeconomic variables X using a two-step Lasso approach:

$$r^e_{t+1} = \beta_0 \mathbb{1}_T + X\beta + \epsilon_{t+1}, \tag{5.11}$$

where $\mathbb{1}_T$ is a length T column vector of ones. All regressors are standardized to have a mean of zero and a variance of one. Our objective is to solve

$$\min_{\beta_0,\beta} \left\{ \|r^e_{t+1} - \beta_0 - X\beta\|_2^2 \right\}, \text{ subject to } \|\beta\|_1 \leq \alpha, \tag{5.12}$$

where α is a (penalty) parameter that determines the degree of regularization. We use cross-validation to select the optimal value for the penalty term that minimizes the mean cross-validated error based on 5,000 simulation runs.

The Lasso method assumes that the coefficients of the linear model are sparse, so the penalty term α sets meaningless small coefficients to zero. This allows us to select only the relevant subset of macroeconomic variables useful for explaining one-month ahead Rolex portfolio excess returns. In a second step, we re-estimate the coefficients on the variables selected from the first step using OLS-regression. The resulting model comprises three macroeconomic time-series which explain 36.80% of variation in one-month ahead Rolex portfolio returns and coefficient estimates are shown in Table 5.10.

Table 5.10: Rolex portfolio returns and macroeconomic variables.

This table shows coefficient estimates of a time-series regression of monthly one-month ahead Rolex portfolio excess returns on a set of selected macroeconomic variables. Relevant macroeconomic variables are determined by a Lasso regression using a total of 126 variables from Federal Reserve Economic Data (FRED) which have non-zero coefficient estimates. All regressors are standardized to have a mean of zero and a variance of one. The sample period is Jan. 2017 - Sep. 2024 (T=98) and $^{*}/^{**}/^{***}$ indicates that coefficients are significantly different from zero at a level of 10%/5%/1% based on Newey and West (1987) corrected standard errors.

Variable	Description	Group	Coefficient	t-statistic
Intercept			0.44^{*}	1.91
COMPAPFFx	3-Month Commercial Paper Minus FEDFUNDS	Interest rate and Exchange Rates	-0.95^{***}	−3.37
AAAFFM	Moody's Aaa Corporate Bond Minus FEDFUNDS	Interest rate and Exchange Rates	0.98^{***}	5.11
EXJPUSx	Japan / U.S. Foreign Exchange	Interest rate and Exchange Rates	-0.55^{***}	−2.60

Rolex portfolio returns are driven by macroeconomic variables only associated with interest and exchange rates. The significance of the Japan / U.S. Foreign Exchange rate is no surprise given that our sample includes watches offered at the Japanese marketplace Rakuten. The net effect of the two interest rate variables is slightly positive, thus Rolex portfolio returns increase when the creditworthiness of the corporate sector is strong (i.e., times when investors are willing to pay a premium to invest in corporate bonds or commercial papers). Moody's Seasoned Aaa Corporate Bond Minus Federal Funds Rate peaked in April 2022 with 3.43% while the Rolex portfolio index was near its all-time high (Feb. 2022). The subsequent downwards movement in Rolex prices is accompanied by the decrease of the Moody's Seasoned Aaa Corporate Bond Minus Federal Funds Rate to an even negative level of -0.21% in July 2024. In general, the influence of interest rates on Rolex returns seems reasonable, given that luxury watches are non-productive assets without any distributions to investors. Similar with gold, interest rates represent an opportunity cost of holding non-productive luxury watches (see O'Conner et al. (2015)).

Typical for luxury goods in general, real personal income, consumption, and variables related with the labor market or price levels (inflation) seem to be irrelevant for luxury watch portfolio returns (see e.g. Aït-Sahalia et al. (2004)). The diversification potential discussed in Section 5.3 is further emphasized by the finding that none of stock market variables included in our set of 126 macroeconomic series is found to be relevant in our Lasso regression.

In summary, the findings in this section reveal that most of luxury watch sellers on the marketplace eBay are professional dealers. Our sample data comprises related ask prices and we detect that returns for luxury watch brand portfolios tend to be lower on Sundays. Because Sunday is typically a day of rest, we conclude that most dealers do not update existing listings on that day, so there is virtually no price discovery resulting in low (ask-price based) returns. In line with our findings and fitting the main characteristics of luxury goods, macroeconomic variables related with the labor market, personal income, inflation, and consumption seem to be irrelevant for one-month ahead Rolex portfolio returns. Instead, returns are mostly driven by macroeconomic variables associated with interest rates and exchange rates.

5.5 Conclusion

The market for luxury watches has experienced rapid growth in recent years. Analyzing this market from the empirical asset pricing point of view is important for multiple reasons: First, the market for luxury watches has so far been completely neglected by academics despite of millions of collectors and investors worldwide using large peer-to-peer platforms like Chrono24 or WatchCharts Marketplace for buying and selling watches. Second, for practitioners, our study offers new insights into the beneficial diversification potential of luxury watches. Interested investors might otherwise avoid participating in the market for luxury watches at all if the risk-return relation remains ambiguous. Third, when implementing the actual watch investment, an investor has to decide when to buy them throughout the seven days of the week and our study helps to answer this question by analyzing potential day-of-the-week effects.

In this study, we are the first to analyze the returns of luxury watches at the brand level over the period January 2017 to September 2024. The brands under investigation are Rolex, Patek Philippe, Audemars Piguet, Cartier, Omega, and Tudor, and related indices generally show an upward movement until the first half year 2022, followed by a decrease in prices until September 2024. An initial \$1 investment in Rolex watches grows to \$1.71, only outperformed by the U.S. stock market (\$2.84) and gold (\$2.30). This corresponds to a weekly return of 0.14%. We find that luxury watch portfolio returns have a remarkably low volatility comparable with U.S. Treasury bonds. For that reason, investments in Audemars Piguet, Patek Philippe, and Rolex watches vastly outperform all our benchmark assets stocks, bonds, and gold on a risk-adjusted basis: Annualized Sharpe ratios for these brands are at least 1.26 and vastly exceed the Sharpe ratio of stocks (0.71) during our sample period. While the luxury watch returns are moderately correlated with each other, they show no co-movement with our benchmark assets, thus indicating beneficial diversification potential. The Covid-19 related stock market turmoils in March 2020 is a good illustration for that potential: While the U.S. stock market accumulated a loss of -27.78% during the three weeks from 03/02/2020 to 03/23/2020, four out of six luxury watch portfolios generated positive returns as high as 1.32% in case of Tudor. Using a dynamic conditional GARCH model (DCC-GARCH) to estimate time-varying correlations, we conclude that luxury watches are indeed a hedge against the stock, bond, and gold market because on average, these returns are unrelated with each other. However, they are at best only a safe haven of very weak nature, because luxury watch returns remain uncorrelated but do not seem to generate positive returns if stocks, bonds, or gold exhibit strong negative returns.

Regarding the optimal timing of an investment in luxury watches throughout the days of the week, we investigate daily return differences using an ARMA-EGARCH(-M) model. Luxury watches can be traded continuously inclusive the weekend and we detect that returns on Sundays are significantly lower than those on other days. This finding is similar with Dorfleitner and Lung (2018) who analyze cryptocurrencies and concludes that lower trading volume as suggested by the mixture of distribution hypothesis (see Tauchen and Pitts (1983)) causes this negative Sundays effect. Our results highlight that non-sophisticated individuals' attention is high on weekends, where investors' mood is near its peak throughout the week. These two effects - trading volume and attention - would, however, imply higher returns on Sundays which contradicts our findings. A possible explanation for the irrelevance of these two economic channels is the peculiar price discovery of luxury watches. In contrast to traditional asset classes where prices on exchanges are determined by the collective actions of buyers and sellers, luxury watches are offered in form of listings on peer-to-peer online marketplaces. In consequence, we observe ask-prices and related returns, reflecting only market opinions from sellers. Most sellers are, however, professional dealers and in line with the fact that Sunday is typically a day of rest in most western countries, we find that listings are only rarely edited on that day. In result, returns for luxury watch brand portfolios are low on Sundays because professional dealers typically do not update offers on Sundays. Testing the relation between an initial set of 126 macroeconomic variables and one-month ahead Rolex portfolio returns, we find that none of macroeconomic variables related with the labor market, personal income, inflation, and consumption seem to be relevant, which fits main characteristics of luxury goods in general. Instead, future Rolex portfolio returns are solely driven by macroeconomic variables associated with interest and exchange rates. This seems plausible given that luxury watches are non-productive assets without any distributions to investors (similar with gold), and interest rates represent opportunity costs.

By understanding the risk-return properties of luxury watches, potential investors can make more informed decisions. In times of an increasing integration of international markets for traditional asset classes with their typically high degree of interconnectedness, luxury watches and collectibles in general may be promising investment opportunities. Identifying the economic factors that drive luxury watch prices and evaluating potential behavioral biases are, however, far beyond the scope of this study. We leave such questions for future research.

Chapter 6

The Global Market for Luxury Watches and Asset Pricing

This research is joint work with Klaus Röder (University of Regensburg). The paper is currently under review in *The Review of Financial Studies*. The journal ranking is *A+* according to the VHB Publication Media Rating 2024. The paper was presented at the 8th Rostock Conference on Service Research 2023 in Rostock, the 30th Annual Meeting of the German Finance Association (DGF) 2024 in Aachen, the 31st Annual Conference of the Multinational Finance Society (MFS) 2025 in Chania, Greece, and is accepted for presentation at the 14th World Finance & Banking Symposium 2025 in Brno, Czech Republic. A previous version of this paper was awarded the *Best Ph.D. Paper Award* by the Multinational Finance Society (MFS) at their 31st Annual Conference.

Abstract

We are the first to analyze the global market for luxury watches between 2010 and 2022 through the lens of asset pricing. We consider a comprehensive list of price- and market-related return predictors in the stock market and construct their watch counterparts using 27,289 watch-month observations. Size, reversal, MAX, and momentum form successful long-short strategies and generate significant return differences. Momentum is inverted and driven by large returns among past losers. Our results favor mispricing related mechanisms based on the analysis of sentiment induced asymmetric pricing effects, cost of arbitrage, and covariance with stock market factors.

Keywords: Empirical Asset Pricing · Alternative Investments · Collectibles · Luxury Watches

JEL classification: D46 · G11 · G12 · G15.

6.1 Introduction

Luxury watches as investment have been completely ignored in top-tier finance and economics journals. This seems surprising to us, as there are at least four key aspects to be considered.

First, wealthy individuals and their decisions shape asset prices to a large extent, and their economic influence seems to be quite large. As a gauge of magnitude, the latest available Deloitte Art & Finance Report 2023 reveals that the investments of ultra-high-net-worth individuals associated with art and collectibles alone already amounts to \$2.174 trillion in 2022.[1] Approximately 63% of surveyed wealth managers have already integrated art into their wealth management offering, and 10.9% of client allocation is associated with art and collectibles in 2023.[2] From a broader perspective, real personal consumption expenditures for jewelry and watches in the USA exceed \$100 billion (measured in 2024-\$, from FRED) in every year since 2021, and the dramatically increase in both desirability and global demand for Rolex watches in recent years is now even discussed in a Harvard Business School case study (Chung (2021)).[3] More important, the economic influence of wealthy households seems to be large enough to shape the markets. Aït-Sahalia et al. (2004) provide evidence that for the very rich, even the equity premium is much less of a puzzle because luxury consumption - in contrast to basic consumption - sufficiently covaries with stock returns to rationalize it. Bali et al. (2023) show that the highly skewed distribution of household wealth explains the anomalous negative risk-return relation among high-volatility stocks,[4] providing evidence that few individual investors affect equilibrium asset prices and can contribute to asset pricing puzzles.[5]

Second, risk and return properties of other collectibles have been broadly studied and have received considerable attention in finance research.[6] Yet, very little is known about luxury

[1] See Deloitte (2023). An ultra-high-net-worth individual is someone with a net worth of at least \$30 million.

[2] Private banks reported an average of 8.6% allocation, and family offices report an average allocation of 13.4% to art and collectibles.

[3] Beyond wealthy households, several studies indicate that between a quarter and a third of all adults in Western countries define themselves as collectors (see Belk (1988) and Pearce (1995)). In a representative survey among 2,000 Germans, Kleine et al. (2024) document that 82.1% of respondents owned collectibles for collecting and/or investment motives at least once in their lifetime. 25.9% of them collect watches, while other areas of collecting are of less interest (with the exception of coins and stamps).

[4] Also known as the idiosyncratic volatility puzzle, see Ang et al. (2006).

[5] Their capability to significantly influence the market is accompanied by the fact that retail investors in general have become increasingly important players in financial markets, see Eaton et al. (2022).

[6] Art (Goetzmann (1993), Campbell (2008), Korteweg et al. (2016)), diamonds (Ariovich (1985), Renneboog and Spaenjers (2012) and Auer and Schuhmacher (2013)), fine wines (Krasker (1979), Dimson

watches. Beyond the main focus of these studies on classical mean-variance analysis in the sense of Markowitz (1952), the broader question arises about whether related investments adhere to these traditional asset-pricing frameworks at all. Recent studies have taken up this challenge by analyzing the sports betting market (Moskowitz (2021)) or cryptocurrencies (Liu et al. (2022)) using standard asset pricing tools. In this meaning, the market for luxury watches is also a useful out-of-sample laboratory that could help to shed some light on asset pricing theories in capital markets, especially since large peer-to-peer marketplaces offer an increasingly easier access to trade luxury watches in recent years.

Third, the rise of peer-to-peer online platforms for luxury watches created ongoing public awareness for the watch market and helped to disseminate that individuals are able to buy and sell luxury watches as convenient as never before.[7] Their key benefit is that they improve liquidity, price discovery, and transaction performance.[8] In other words, these platforms provide the much-needed transparency for luxury watch investors, which further attracts new collectors and investors. Google Trends reveals that the global popularity of the search term *"Rolex investment"* steadily rose since the beginning of our sample period in 2010 and peaked in February 2022.[9] Moreover, luxury watches also receive wide attention among social media. As of October 2025, the official channel of Rolex has 2.47m subscribers on the popular video platform YouTube. Together with similar channels run by retailers like *Teddy Baldassarre* (1.33m subscribers) or ones that are highly specialized on luxury watches, e.g., *Watchfinder & Co.* (1.16m subscribers), *Wristwatch Revival* (1.07m subscribers), and many others, they account for billions of views.

Fourth, the growing popularity for luxury watches as investments gradually leads to a growing market efficiency, and our data supports this view. For instance, Patek Philippe announced in January 2021 that their popular model Nautilus 5711 (approx. $35,000

et al. (2015), see Le Fur and Outreville (2019) for a review of the literature) and cars (Martin (2016), Laurs and Renneboog (2019), Le Fur (2023)), as well as less common collectibles such as antique furniture (Rush (1968)), coins (Kane (1984)), timber (Redmond and Cubbage (1988)), antique firearms (Avery and Colonna (1987)), Stradivarius violins (Ross and Zondervan (1989)), photographs (Pompe (1996)), stuffed animals (Burton and Jacobsen (1999)), sculptures (Locatelli-Biey and Zanola (2002)), postage stamps (Dimson and Spaenjers (2011)), whisky (Moroz and Pecchioli (2019)), non-fungible tokens (Dowling (2022)), LEGO sets (Dobrynskaya and Kishilova (2022)), Magic the Gathering game cards (Langelett and Wang (2023)), comics (Bocart et al. (2023)), and video game skins (Reichenbach (2025)).

[7]A New York Times article acknowledged that the rise of online platforms for trading luxury watches, the growing influence of social media, and the first widespread effects of the best-selling Apple Watch since 2015 fueled demand for mechanical luxury watches (see Gomelsky (2023)).

[8]These platforms typically have very high standards for sellers. See Section 6.2.1 for further details.

[9]The consistent increase in popularity is also observed for similar search terms like *"is a Rolex a good investment"*, *"Rolex submariner"*, and *"Rolex price"*.

MSRP) will be discontinued. The average market price for that model in our sample increased from \$84,324 in December 2020 to \$121,958 in February 2021, a surge of 45% within two months. Andrioli (2021) provides more detailed evidence for the short-term efficiency of the luxury watch market: First rumors on the announcement of discontinuation spread in online forums on January 22, 2021. Average prices on peer-to-peer marketplaces for luxury watches immediately increased by 25% on that day. When Patek Philippe officially released the discontinuation three days later, prices increased by an additional 4.8%. Given these immediate price reactions, we suggest that the cross-section of luxury watch returns can be meaningfully analyzed using standard asset pricing methods.

We hand-collect monthly prices for 345 luxury watches of 20 brands from the world's largest peer-to-peer marketplace for luxury watches Chrono24.com (`https://www.chrono24.com`). Our final sample consists of 27,289 watch-month observations between June 2010 and March 2022. In this regard, the sample opens new possibilities to test theories of cross-sectional asset pricing anomalies. We test 30 characteristics related with size, value, momentum, and volatility effects and find that size (AGE, number of month listed on Chrono24.com), reversal (past one-month return), short-term momentum (past four-to-one month return), and MAX (Bali et al. (2011)) generate significant difference returns among zero-investment quintile portfolio strategies. Both the k-FWER test method by Lehmann and Romano (2005) and an F-test for the joint significance provide evidence that our results are unlikely to generate by chance.

Size (AGE), i.e., long high size (AGE) and short low size (AGE) to be specific, generates a highly significant *positive* premium of 0.48% per month until February 2020 and we provide evidence that the effect is not related with attention which we proxy by Google search results at the watch level. Looking at absolute prices (PRC) as a measure for size (Miller and Scholes (1982)), we find that least expensive watches have higher idiosyncratic volatility and higher standard deviations of past returns, similar to microcaps. Controlling for cost of arbitrage as in Stambaugh et al. (2015), the size (PRC) effect is only prevalent among high cost of arbitrage watches with a premium of highly significantly -1.04%.

Momentum strategies reveal that past losers generate highly significant and economically large *positive* returns for measurement periods up to past 16 months, among all price levels except for most expensive watches. For holding periods up to seven months after the initial portfolio sort, we reject the null hypothesis of difference returns being zero at the 99% confidence level. Negative returns of past winners revert towards zero, but past losers remain above the level of their initial response for all subsequent holding periods.

In favor of an underreaction based momentum mechanism, difference returns of high-low momentum strategies using past four-to-one month returns are a highly significant -1.57% for small watches and -0.58% for the mid-size group. An initial overreaction that is linked with investor attention as in Peng and Xiong (2006) seems to be unlikely as momentum strategies remain highly significant even among low-attention watches.

In accordance with the findings of Stambaugh et al. (2015) on the asymmetric pricing effect of sentiment, we find that sentiment-related variation in the performance of size, reversal, momentum and MAX strategies is mainly due to their short positions. Our results are in favor of a mispricing related interpretation and that the strategies reflect a mispricing commonality across luxury watches. We use factor returns for the models proposed in Fama and French (2015), Fama and French (2018), Daniel et al. (2020a) and Stambaugh and Yuan (2017) to account for possible covariation of our portfolio strategy returns across different asset classes as shown in Asness et al. (2013). Our factor-model adjusted alphas of luxury watch investment strategies are quantitatively similar to our portfolio results. We detect that reversal tends to negatively comove with the factor FIN that captures predominantly longer-term mispricing and correction. Similarly, momentum is weakly related with the factor PERF, supporting our mispricing related interpretation.

The rest of the paper is organized as follows. Section 6.2 describes the data. Section 6.3 examines counterparts to well established equity return predictors for the cross-section of luxury watch returns. Section 6.4 provides additional results and robustness tests. Finally, Section 6.5 concludes.

6.2 Data and variables

6.2.1 Luxury watch data

The world's largest[10] peer-to-peer marketplace for luxury watches is Chrono24.com (`https://www.chrono24.com`). Founded 2003 in Germany, approximately 3,500 commercial dealers (mostly jeweler's) from over 100 countries offer regularly more than 500,000 luxury watches, comprising an aggregated value of roughly 6 billion Euro.[11]

Chrono24.com requires that professional dealers must provide their photo ID, commercial register entry, business address, and tax number, before they can sell watches on their platform. Private sellers must pass an inspection before they publish their offers. As part

[10]Based on own company statements.
[11]Although private sellers are permitted, more than 90% of all sellers are professional dealers.

of the inspection, the seller must provide additional proof of ownership in the form of two pictures of the watch set to specific times. Further, they offer an escrow service that assures a secure transaction process for both buyers and sellers, guarantee of authenticity, compulsory insured shipments, and a 14-day return policy upon delivery. These obligations clearly help to prevent listings of dubious or implausible offers and results in more reliable price data.

After registering on their platform and signing in to the account, Chrono24.com displays its users charts with monthly price information for most popular offered luxury watches.[12] An exemplary chart for a Patek Philippe Nautilus (reference no. 5711/1A-010) is shown in Figure 6.1.

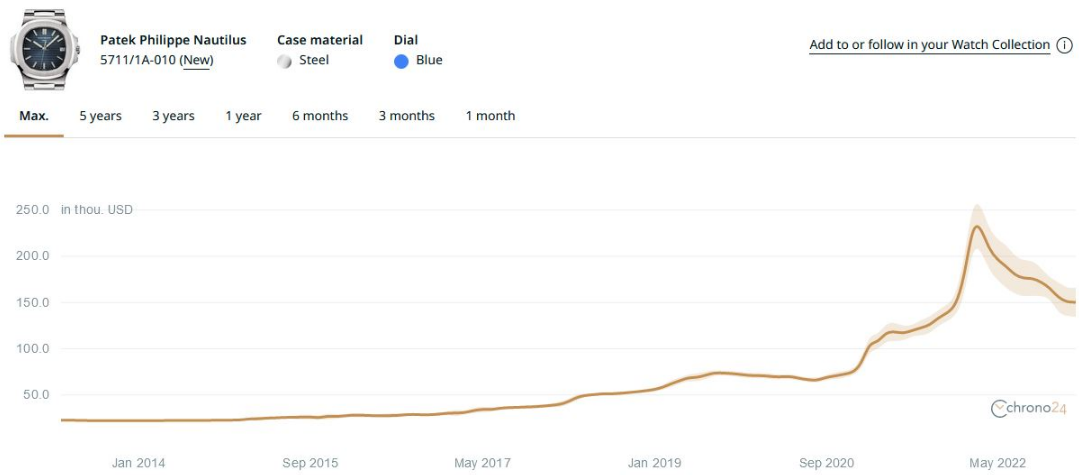

Fig. 6.1. Exemplary chart with monthly frequency (only displayed for registered users) from Chrono24.com of a Patek Philippe Nautilus, reference number 5711/1A-010. The graph displays (monthly) time-series means of average daily ask prices across all worldwide offers for that watch model within a month. The shaded area covers the according time-series average of daily cross-sectional maxima and minima prices.

It is necessary to sound a note of caution on the interpretation of these price data. In several discussions with Chrono24.com, they have disclosed the most important facts on how they process price information but refused to go into further details than discussed hereinafter.

First, price charts are only shown for popular watch models, i.e., when the amount of according worldwide offers is sufficiently large. It is important to know, that the name of a model itself is not the unique identifier of a specific watch, but rather its reference number. Take e.g., the Rolex GMT-Master II: We have six distinct time-series price data for that model in our sample, because they differ e.g., in the color of their dial or bezel,

[12]Creating a user account only requires a valid email address and password.

the material or size of the case, the movement or their general condition (i.e., if the watch is new/unworn or vintage). These characteristics are differently assessed by collectors and investors. While the price for a GMT-Master II ref. 126710BLNR (nicknamed *Batgirl*) with black dial and blue/black bezel has an average ask price of $27,252 in March 2022, the GMT-Master II ref. 116710LN also having a black dial, but a bezel entirely in black, costs only $17,538.

Second, price data is based on available offers and thus refers to ask prices. Shipping costs, taxes, customs, and possible other charges related with import duties are not incorporated.

Third, price data is available with monthly frequency back to June 2009, whereas daily prices are only tracked for the most recent month. The main graphs, as shown exemplary in Figure 6.1, display monthly time-series averages of (non-disclosed) daily cross-sectional mean prices across all worldwide offers for that watch model reference within a month. The shaded area covers the according time-series mean of daily cross-sectional maxima and minima prices. In other words, each day, Chrono24.com calculates the mean, minimum, and maximum price across all worldwide offers on their platform, but do not disclose these figures. Then, for each calendar month, they report related (time-series) averages of these non-disclosed daily prices. We address this issue in more detail in Section 6.2.2 and Section A.I of the Appendix.

Fourth, Chrono24.com started to track ask prices in June 2009, but these price charts are shown registered users only recently. Given the fact that they are a company located in Germany, they convert all historic prices of non-Euro listings into currency Euro, but using the exchange rate at the time the price chart is retrieved. It is, however, easy to avoid this issue of implicitly static exchange rates: We retrieve all prices in Euro and then convert the according time-series to U.S.-$ using exchange rates from LSEG Eikon.[13]

Last, while buying a luxury watch on Chrono24.com is feasible at any time, we have reasonable evidence that selling luxury watches is possible quite fast, so the market for luxury watches is somehow liquid.[14] Based on WatchCharts Analytics, a research

[13]Recall that monthly watch prices are actually means of daily prices within a month. We convert them using end of month exchange rates. In unreported calculations using either mid of month or daily averages of exchange rates, we conclude that our results are not affected by that decision.

[14]Of course, luxury watches are far from being traded as liquid as stocks, but two key properties should be noticed: First, for a given reference number, watches are interchangeable and fungible. Second, Chrono24.com is a large peer-to-peer platforms, so we observe periodically sales in stark contrast to occasional and infrequent auctions conducted by auctioneers. Based on own corporate statements, the transaction volume generated already in 2017 exceeded one billion Euro (see Maillard (2018)).

platform collecting and structuring millions of data points for over 100 brands and 25,000 watches, the median days a watch is listed on one of their analyzed marketplaces are no longer than one month, typically between 15 and 25 days. This is, however, based on only few selected observations for watches in our sample. Unfortunately, Chrono24.com does not disclose information on the metric "median days on the market" for its marketplace and WatchCharts Analytics observes platforms like eBay, Rakuten Japan, and Carousell, but not Chrono24.com. Based on own evidence by looking at selected watch listings on Chrono24.com using bookmarks, we confirm that most listings are not available after two weeks. Nevertheless, frequently trading luxury watches is feasible through the lens of asset pricing. Further evidence can also be derived from the fact that Chrono24.com requires sellers to have at least 50 watches to sell with an average value of $2,000 each to be classified as a professional dealer with related listing packages for selling up to 1,000 watches a month (for a fixed fee of 2,199 Euro).[15]

We collect monthly mean, minimum, and maximum ask prices for 345 watches from 20 brands between June 2009 and March 2022. Because the first year of this sample comprises only 27 watches, we begin our analysis in June 2010, observing at least 56 watches. In total, the sample consists of 27,289 watch-month observations.

The market for luxury watches is dominated by companies located in Switzerland. According to Morgan Stanley's annual watch report (see Morgan Stanley and LuxeConsult (2024)), leading Swiss manufactures in terms of worldwide retail market share 2024 are Rolex (32%, together with their brand Tudor), Swatch Group (18%, among Omega, Longines, Blancpain and others), Richemont(19%, Cartier, Jaeger-LeCoultre, IWC, Piaget), Patek Philippe (7%) and Audemars Piguet (5%).[16] This is reflected in our sample of luxury watches presented in Panel A of Table 6.1. 52 watches in our sample are produced by Rolex, followed by 37 Omega and 35 Patek Philippe watches. The latter accounts on average for 35.28% of aggregated watch prices, indicating that these watches are among the most expensive ones in our sample. This is confirmed in Panel B where distributional characteristics for watch prices are presented. A watch made by Patek Philippe has an average price of $51,964. On average, we observe 192 watches per month with a price of $13,614. Prices are positively skewed, and their 5th (95th) percentile price is $2,107 ($48,418). The last two rows in Panel B shows summary statistics for prices belonging to

[15]See Section 6.4.7, which explicitly calculates luxury watch strategy returns net of transaction costs, for further information.

[16]Most of them are very secretive about the number of produced and sold units, but the Federation of the Swiss Watch Industry reports wristwatch export figures (including non-luxury watches) to be a total of CHF 25.993b in 2024 with most of the exports going to the USA (17%), China (8%), and Japan (8%).

Table 6.1: Summary statistics of luxury watches.

This table shows summary statistics for the brands included in our sample of luxury watches and their ask price from the world's largest peer-to-peer marketplace for luxury watches, Chrono24.com. Panel A reports the total number of watch models for each brand and the brand's average percentage weight in terms of aggregated prices. Each model is identified by a unique reference number for that type of watch, having specific characteristics e.g., the color of dial and bezel, the size and material of the case, and its classification as being vintage or new/unworn. Panel B reports monthly time-series averages of the cross-sectional mean, standard deviation (SD), 5th percentile, 25th percentile, median, 75th percentile, 95th percentile, and the number of observations (N) of global ask prices for luxury watches. The sample period is June 2010 to March 2022 and all data is denominated in U.S.-$.

Panel A: Luxury watch brands

Brand	Watches	Weight	Brand	Watches	Weight
A. Lange & Söhne	8	4.61	IWC	18	3.93
Audemars Piguet	16	8.22	Jaeger-LeCoultre	16	4.25
Blancpain	11	2.25	Omega	37	3.57
Breitling	18	1.82	Panerai	13	2.39
Cartier	15	2.07	Patek Philippe	35	35.28
Chopard	15	2.65	Piaget	7	0.84
Franck Muller	11	2.00	Rolex	52	15.96
Girard Perregaux	10	0.83	TAG Heuer	13	1.01
Glashütte Original	7	3.10	Tudor	15	1.33
Hublot	16	2.56	Zenith	12	1.34

Panel B: Luxury watch prices (in U.S.-$)

Watches	Mean	SD	5th	25th	Median	75th	95th	N
All	13,614	22,049	2,107	4,024	6,758	12,701	48,418	192
Rolex	10,532	7,269	3,279	5,475	9,072	11,767	26,530	34
Patek Philippe	51,964	45,964	13,828	25,794	41,416	60,581	108,864	20
Audemars Piguet	33,085	27,812	17,645	20,778	24,480	33,470	75,919	8
Other	7,620	8,391	1,944	3,566	5,247	8,533	20,495	131
Low ($< P_{20\%}$)	2,602	682	1,548	2,102	2,581	3,207	3,543	39
High ($> P_{80\%}$)	42,586	36,134	17,827	22,206	29,990	48,700	91,644	39

the bottom, resp. top, quintile. We observe that least expensive watches have a relatively small standard deviation of $682, resp. 26.21% of their average price, compared to most expensive ones ($36,134, resp. 84.85%).[17]

6.2.2 Econometric considerations and summary statistics

Before we turn to our main analysis on cross-sectional expected returns, we have to address an important concern related with our price data for luxury watches: As mentioned in the previous section, we do not observe end-of-month data which are most common in asset pricing studies. Instead, our monthly prices represent averages of non-observable daily prices within a given month (we refer to this approach as mean-of-month prices). Intuitively, if daily prices close to the end of a month suddenly drop compared to previous prices that month so far, returns are overestimated and vice versa.

To provide an example, Figure 6.2 shows the S&P 500 Index during Covid-19 related market turmoils in the first half of 2020, and we replicate the mean-of-month methodology from Chrono24.com as well as the common end-of-month approach for comparison. The index closed at 3,226 points on January 31, 2020, and the average index level was 3,278 points that month. Until February 19, the index has increased to 3,386 points but plunged within the last seven trading days that month to 2,954 points. Based on end-of-month index levels, its index return was -8.41%. However, the average index level in February was 3,282 points, resulting in a monthly return of 0.12%. This is because the positive return of 4.98% until February 19 accounts for roughly two thirds of trading days and offsets the -12.76% the last week. On the other hand, the bias in monthly returns tends to mean-revert in the following month mechanically: Based on the lower index level of 2,954 points in February 28 compared to Februaries average of 3,282, end-of-month return in March 2020 were -12.51% while the Chrono24.com methodology results in a now lower return of -19.19%.

To evaluate the magnitude of this potential bias in our watch returns, we rely on 430 equity portfolios for 47 anomalies described in Haddad et al. (2020), that have non-missing daily and monthly returns from July 2010 to March 2022 (our sample period). First, we recall the mean-of-month procedure of Chrono24.com by converting daily returns for each

[17] Our data from Chrono24.com reflects market prices formed by demand and supply which greatly differ from MSRP's. The retail price of a Rolex Daytona (reference 116520) is approximately $12,500, depending on location and taxes, whereas the average ask price on Chrono24.com in March 2022 is $46,643.40. However, historical MSRP's are generally not available due to the secretiveness of the Swiss watch manufactures. Noteworthy, the MSRP (minus a discount for counterparty risk etc.) marks a latent minimum ask price under the assumption of rational sellers.

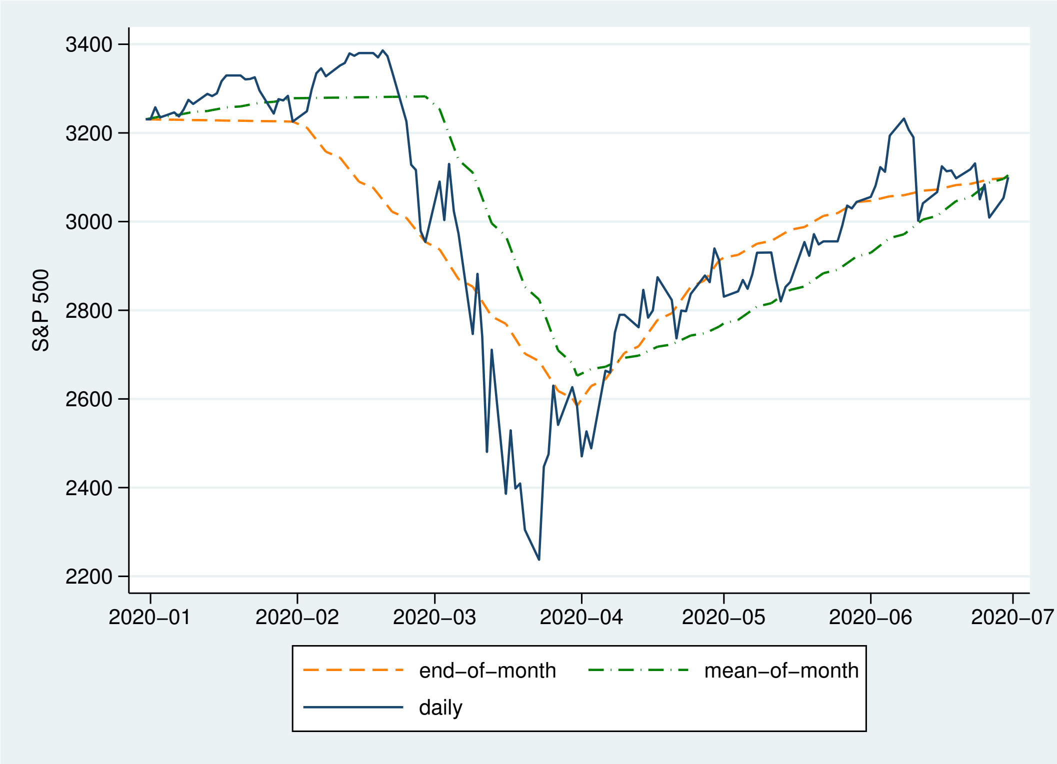

Fig. 6.2. This figure provides a visual illustration of the difference between common end-of-month return calculation and the mean-of-month approach used by Chrono24.com. The solid line (blue) shows the S&P 500 index for the first half of 2020 at the daily level. The dashed line (orange) are related monthly returns (cumulated) using end-of-month index levels. The dash-dotted line (green) refers to the mean-of-month approach used by Chrono24.com, where averages of daily index levels within a given month are used to calculate monthly (cumulative) returns. Monthly cumulative returns are linearly interpolated within each month for display purposes accounting for mixed time frequency (daily and monthly).

portfolio i into a series of prices, normalized to $1 at the end of June 2010. Each month, we calculate the average price of all daily prices and denote this as our monthly price $P_{i,t}^{mean}$. Associated monthly returns are obtained by $r_{i,t}^{mean} = (p_{i,t}^{mean}/p_{i,t-1}^{mean}) - 1$. Finally, we compare the distribution of mean-of-month derived returns $r_{i,t}^{mean}$ with common end-of-month calculated returns $r_{i,t}$.

Among the 430 anomaly portfolios, we observe an average end-of-month return of 1.26%. A Welch test on the difference of $r_{i,t}^{mean} - r_{i,t}$ reveals a highly significant mean of -0.0080% (t-statistics: -12.48). Without dispute, we leave this slightly underperformance of mean-of-month returns as economically negligible. On the other side, the time-series volatility of monthly mean-of-month based returns is underestimated by 0.74 percentage points on average. Given an average end-of-month based volatility of 3.82% p.m., our estimated mean-of-month volatility is off by more than 20% (i.e., 0.74/3.82). For a better understanding of this bias, we estimate a pooled regression $r_{i,t} = \alpha_i + \beta_i r_{i,t}^{mean} + \epsilon_{i,t}$, in which the dependent variable is a testportfolios' monthly return based on end-of-month calculations, and the explanatory variable is our mean-of-month equivalent using the average of daily index levels within a month for return calculations. If mean-of-month returns fully reflect the information in end-of-month returns, we would observe $\alpha = 0$ and $\beta = 1$. While α is indeed an insignificant 0.2304, β is 0.8219 with a standard error of 0.0762 (clustered by both portfolio and month). Using the representation $\hat{\beta} = \rho \cdot \frac{\sigma_{r_{i,t}}}{\sigma_{r_{i,t}^{mean}}}$, we observe a strong correlation $\rho = 0.66$ and a ratio of volatilities of 1.24. Further, the R^2 of the regression is a modest 44.23%. In conclusion, these results indicate that using mean-of-month based returns as provided by Chrono24.com results in valid point estimates for average returns, while related volatilities are strongly downwards biased compared to common end-of-month returns.

To account for the volatility bias and to err on the side of caution, we decide to scale volatilities for each series i of luxury watch returns $r_{i,t}^w$ by a factor of 1.31, i.e., we apply the transformation $\bar{r}_{i,t}^w + 1.31 \cdot \left(r_{i,t}^w - \bar{r}_{i,t}^w\right)$ with $\bar{r}^w$ as the vector of demeaned returns.[18] Although this choice is somehow arbitrary, this scaling factor corresponds to the 95th percentile of estimated volatility ratios $\frac{\sigma_{r_{i,t}}}{\sigma_{r_{i,t}^{mean}}}$ (see Section A.I of the Appendix) and seems to be rather conservative. The main advantage of this simple transformation is that this setup avoids reliance of our results on further assumptions on the unobservable daily watch returns needed in more advanced techniques like ARMA- or GARCH-models.[19]

This can be illustrated briefly by applying this volatility scaling approach to the U.S. equity market return (July 2010 - March 2022):[20] Average end-of-month based market returns are 1.30% p.m. compared with 1.25% p.m. following the mean-of-month approach, and the modest return difference is statistically insignificant. Prior to scaling, the monthly

[18] If our estimates for monthly luxury watch return volatilities are consistently too high, reported Sharpe ratios and statistical inferences for our analyzed trading strategies may be too conservative, respectively, rendered insignificant.

[19] We face the same problematic as the literature on real estate pricing which concurs that returns calculated from appraisal-based real estate indices incorrectly reflect and underestimate the true real estate's volatility (see Goetzmann (1992) and Maurer et al. (2004)). To recover unbiased estimates, Blundell and Ward (1987) and Firstenberg et al. (1988) suggest the meanwhile common transformation of the residuals from an auto-regressive process that was fitted to the original series. The choice of an appropriate unsmoothing procedure is, however, problematic, due to the assumptions about the appraisal process, index construction process and market inefficiencies (see Lai and Wang (1998)). Similarly, another strand of literature addresses a related problem, namely, how to unsmooth returns of illiquid assets held by funds and traded only infrequently (see Couts et al. (2024)). For robustness of our simple approach and for comparison, Section A.I of the Appendix reports results from applying related ARMA-models to unsmooth luxury watch returns to recover unbiased volatility estimates.

[20] Data is provided by Kenneth French data library.

volatility of mean-of-month returns is 3.35%, and scaling by 1.31 results in a volatility estimate of 4.39%. The true end-of-month based volatility of market returns is 4.11% and the difference of 0.28 percentage points is statistically insignificant (F-statistic: 1.13). For this reason, our procedure of scaling volatilities for each series of luxury watch returns seems applicable and reasonable.

We construct a luxury watch market return $Watch^{Dow}$ as the price-weighted return of all underlying available watches. Economically, this allows investors to replicate our portfolio strategy in a quite reasonable way for watch collectors, by buying exactly one piece of each watch model and thus following a Dow-Jones like strategy. Similarly, we construct indices at the brand level for Rolex, Patek Philippe, and Audemars Piguet, reflecting their economically importance among the market for luxury watches.

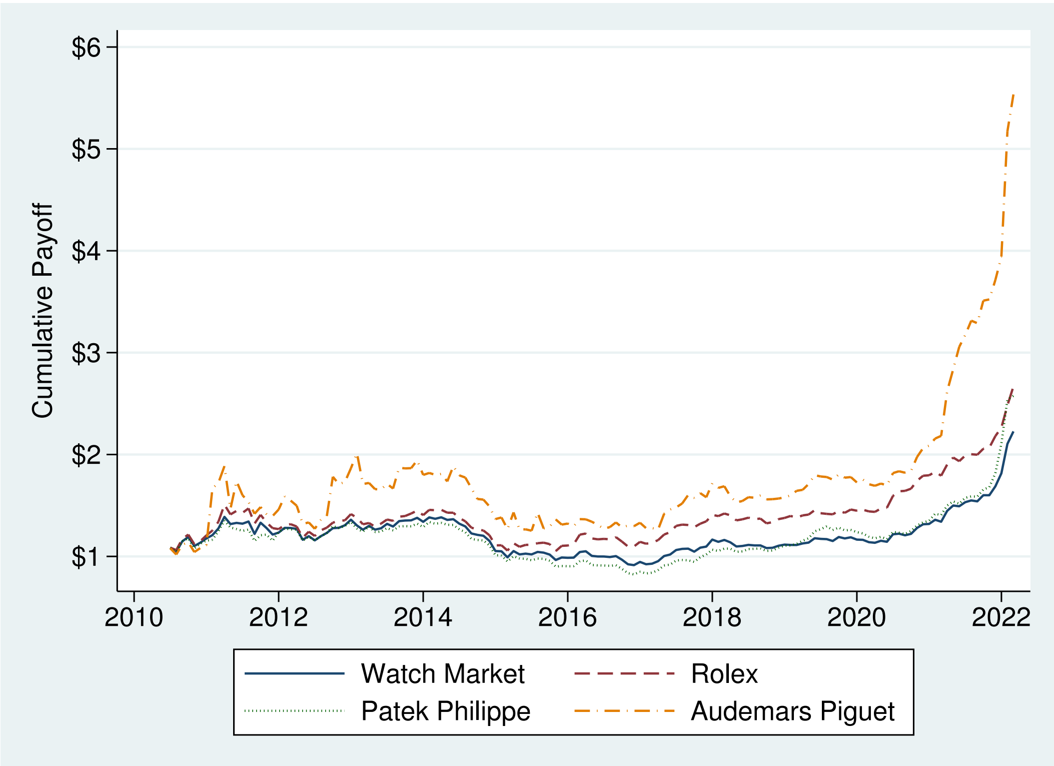

Fig. 6.3. This figure illustrates the cumulative payoff of a $1 investment in the portfolio of price-weighted luxury watches. The market portfolio comprises all luxury watches in our sample, whereas the brand indices only include watches from a specific brand. The sample period is July 2010 to March 2022.

Fig. 6.3 illustrates the cumulative payoff of a $1 investment in the portfolio of price-weighted luxury watches at the market level and brand level for Rolex, Patek Philippe, and Audemars Piguet. The initial investment of $1 in the luxury watch market portfolio at the end of June 2010 results in a terminal value of $2.23 until March 2022.[21] We observe that prices for luxury watches strongly increased at the beginning of 2020. Looking at Audemars Piguet, the price-weighted index generated $1.71 until March 2020 and subsequently surged to $5.53 until March 2022, which implies an average return of 5.02% p.m., resp., 60.21% annualized, within a two-year period. This is, however, not driven by a single watch. Although Audemars Piguet is mainly known for its model Royal Oak, all variants in our sample (Self-winding, Chronograph, Jumbo, Offshore, etc.) at least doubled in prices.[22] However, in a mixture of Covid-19 related restrictions resulting in a surge of household savings, essentially zero interest rates and zero inflation in most developed countries around the world, we observe similar price increases in other alternative markets like cryptocurrencies, non-fungible tokens, or collectibles like art (old masters and contemporary art).[23] In addition, the U.S. government enacted fiscal stimulus during that time distributed approximately $814 billion to taxpayers and a significant portion of that was used for investments (see Greenwood et al. (2023)).

Related monthly returns for the performance of luxury watches are presented in Table 6.2. Panel A presents results for the whole sample period of July 2010 to March 2022. The price-weighted market index of luxury watches $Watch^{Dow}$ has an average return of 0.64% per month and none of our return series shows significant one-month autocorrelation. Taking into account related risk (i.e., volatility), the annualized Sharpe ratio of price-weighted luxury watch indices is within the range 0.55 to 0.68. There is, however, a huge discrepancy between the performance until the beginning of 2020 and afterwards. Panel B and C show average returns for both sub-periods and it is apparent that luxury watches showed a remarkable performance since 2020. The order of outperformance is huge as indicated by an annualized Sharpe ratio of 2.33 for the entire watch market or even 2.79 for Rolex models in the latter period. However, from July 2010 to February 2020, luxury watches yielded a moderate return of 0.19% at the market level. For this reason, our analysis on cross-sectional predictors of luxury watch returns in the next section consistently reports results for both sub-periods.

[21]This is considerably more than a likewise investment in international governmental bonds (FTSE World Government Bonds Index) or commodities (S&P GSCI Index), which results in a total value of $1.16, resp. $0.92, over the sample period. However, a global developed countries stock market portfolio generates $3.67. For comparison with other collectibles, Section A.II of the Appendix shows related returns for fine wine and art.

[22]Take e.g., the Royal Oak Jumbo Ultra Thin, reference number 15202ST.OO.1240ST.01. On average, the ask price was $46,317.12 in March 2020 and $173,362.00 in March 2022.

[23]Bitcoin price increased from $6,437.31 in 04/01/2020 to $45,554.16 in 04/01/2022 (data from CoinMarketCap).

Table 6.2: Summary statistics for the time-series of monthly luxury watch returns.

This table presents summary statistics for the monthly time-series of luxury watch returns. $Watch^{Dow}$ represents the price-weighted market return comprising all luxury watches in our sample. Similarly, $Watch^{Rolex}$, $Watch^{Patek\ Ph.}$, and $Watch^{AP}$ denote price-weighted indices for luxury watches from Rolex, Patek Philippe, and Audemars Piguet. The table shows the mean, standard deviation (SD), minimum (Min), median, maximum (Max), skewness (Skew), excess kurtosis (Kurt), annualized Sharpe ratio (SR), and the first order autocorrelation coefficient (AC(1)) for the time-series of returns.

Panel A: Full sample - July 2010 to March 2022									
	Mean	SD	Min	Median	Max	Skew	Kurt	SR	AC(1)
$Watch^{Dow}$	0.64	3.74	−9.06	0.42	15.90	0.42	4.65	0.55	0.01
$Watch^{Rolex}$	0.77	3.70	−10.11	0.67	12.52	0.07	3.87	0.68	−0.03
$Watch^{Patek\ Ph.}$	0.75	4.11	−8.87	0.44	20.54	0.96	7.08	0.60	0.14
$Watch^{AP}$	1.50	7.90	−22.28	0.79	49.98	2.12	14.11	0.64	−0.08
Panel B: Subperiod - July 2010 to February 2020									
	Mean	SD	Min	Median	Max	Skew	Kurt	SR	AC(1)
$Watch^{Dow}$	0.19	3.54	−9.06	0.18	9.79	0.03	3.49	0.14	−0.15
$Watch^{Rolex}$	0.39	3.73	−10.11	0.48	12.52	0.09	3.97	0.32	−0.11
$Watch^{Patek\ Ph.}$	0.24	3.60	−8.87	0.22	9.71	−0.10	3.35	0.19	−0.08
$Watch^{AP}$	0.77	7.87	−22.28	0.30	49.98	2.28	16.15	0.32	−0.18
Panel C: Subperiod - March 2020 to March 2022									
	Mean	SD	Min	Median	Max	Skew	Kurt	SR	AC(1)
$Watch^{Dow}$	2.71	4.01	−1.97	1.75	15.90	1.47	5.63	2.33	0.26
$Watch^{Rolex}$	2.49	3.08	−2.09	2.33	9.61	0.62	2.68	2.79	0.14
$Watch^{Patek\ Ph.}$	3.14	5.40	−2.72	1.50	20.54	1.84	6.29	2.01	0.37
$Watch^{AP}$	4.92	7.21	−2.59	3.85	31.19	2.28	8.65	2.36	0.15

6.3 Cross-sectional predictors of luxury watch returns

Many investment strategies in the equity market are adapted to other asset markets to establish stylized facts for theoretical models. Asness et al. (2013) document value and momentum effects in currency, commodity and bond markets and Liu et al. (2022) find size and momentum effects among cryptocurrencies. We follow Liu et al. (2022) and consider a comprehensive list of well-established equity strategies from Feng et al. (2020) and Chen and Zimmermann (2022) and select characteristics that can be constructed using our price information of luxury watches. This procedure ensures that we do not choose ex-ante cross-sectional watch return predictors that suit our analysis. We group a total of 30 characteristics into four categories (size, value, momentum, and volatility) and provide a summary of our results in Table 6.3.

For the remainder of this section, we analyze zero-investment strategies based on each characteristic. At the end of each month, we sort luxury watches into quintile portfolios based on the value of the respective characteristic. We calculate price-weighted portfolio returns over the risk-free rate for the subsequent month. Price-weighted returns ensure an easily replicable investment strategy of buying one of each watch models in a certain portfolio. We finally calculate excess returns of the long-short strategy as the difference between the fifth and the first quintile portfolio. Shorting luxury watches is hardly feasible, so our investment strategy is not only judged by its long-short hedge return, but also by being long-only the first or fifth quintile portfolio, financed by the risk-free rate.[24]

[24] Luxury watch brands are using blockchain technology to provide proof of ownership and authenticity of watches through digital certificates, resp. non-fungible tokens (NFTs). Since 2020, all Breitling watches come with a blockchain-based digital passport and Audemars Piguet and Vacheron Constantin announced similar plans. Beyer Chronometrie (a watch store located in Zurich, Switzerland) even launched non-fungible watches (NFWs), i.e., a fully digital watch stored on the blockchain and not linked to any physical item. Investors can buy fractional units of watches in NFT/NFW sales on platforms like OpenSea or Chronology.io. Similar to short-selling, investors receive rewards in form of cryptocurrency tokens if they stake, i.e., lock up, these digital assets for a predefined period of time. See Galanti (2022) for further information.

Table 6.3: Characteristics and predictors of cross-sectional luxury watch returns.

This table provides definitions and references for well-established characteristics from Feng et al. (2020) and Chen and Zimmermann (2022) that predict the cross-section of stock returns. We analyze their according counterparts in the luxury watch market. Panel A lists predictors with significant (at the 5% level for the full sample period) zero-investment price-weighted watch portfolio returns based on monthly quintile sorted watch portfolios. Insignificant predictors are shown in Panel B.

Category	Predictor	Return	t-statistic	Reference	Definition
Panel A: Significant Predictors					
Size	AGE	0.34	2.27	Barry and Brown (1984)	Number of month listed on Chrono24.com.
Mom	r 1,0	-0.73	-1.41[a]	Jegadeesh and Titman (1993)	Past one-month return.
Mom	r 2,0	-1.34	-2.68	Jegadeesh and Titman (1993)	Past two-month return.
Mom	r 3,0	-1.30	-2.61	Jegadeesh and Titman (1993)	Past three-month return.
Mom	r 4,0	-1.33	-2.50	Jegadeesh and Titman (1993)	Past four-month return.
Mom	r 4,1	-0.21	-0.60[b]	Jegadeesh and Titman (1993)	Past one-to-four-month return.
Mom	r 16,0	-0.89	-1.98	Jegadeesh and Titman (1993)	Past 16-month return.
Vol	MAX	-0.65	-2.30	Bali et al. (2011)	Maximum return in the portfolio formation month. Measured as the ratio of max. price in t_0 and min. price in t_{-1}.
Panel B: Insignificant Predictors					
Size	PRC	0.11	0.29	Miller and Scholes (1982)	Log price in the portfolio formation month.
Size	MAXDPRC	0.14	0.34	George and Hwang (2004)	Maximum price in the portfolio formation month.
Value	VAL_AMP	-0.67	-1.17	Asness et al. (2013)	Log of the average price from 4.5 to 5.5 years ago, divided by the most recent price.
Value	VAL_FF	-0.72	-1.42	Fama and French (1996)	Negative of the past 13-to-60-month return.
Mom	r 8,0	-1.07	-1.91	Jegadeesh and Titman (1993)	Past eight-month return.
Mom	r 12,1	-0.01	-0.03	Carhart (1997)	Past twelve-month return skipping the most recent one.
Vol	BETA1Y	-0.05	-0.18	Fama and MacBeth (1973)	The regression coefficient β^i in $r_i - r_f = \alpha_i + \beta^i \left(Watch^{Dow} - r_f\right) + \epsilon_i$. The model is estimated using past 12 months exc. returns req. at least 10 observations.
Vol	BETA2Y	0.06	0.17	Fama and MacBeth (1973)	BETA1Y but using monthly returns over the past 24 months req. at least 20 obs.
Vol	BETA3Y	0.27	0.65	Fama and MacBeth (1973)	BETA1Y but using monthly returns over the past 36 months req. at least 24 obs.
Vol	BETA1Y^2	0.05	0.16	Fama and MacBeth (1973)	BETA1Y squared.
Vol	BETA2Y^2	0.07	0.18	Fama and MacBeth (1973)	BETA2Y squared.
Vol	BETA3Y^2	0.23	0.57	Fama and MacBeth (1973)	BETA3Y squared.
Vol	IVOL1Y	-0.48	-1.70	Ang et al. (2006)	Idiosyncratic volatility, measured as the standard deviation of the residual after estimating $r_i - r_f = \alpha_i + \beta^i \left(Watch^{Dow} - r_f\right) + \epsilon_i$. The model uses past 12 months returns and requires at least 10 obs.
Vol	IVOL2Y	-0.18	-0.70	Ang et al. (2006)	IVOL1Y but using monthly returns over the past 24 months req. at least 20 obs.
Vol	IVOL3Y	-0.01	-0.03	Ang et al. (2006)	IVOL1Y but using monthly returns over the past 36 months req. at least 24 obs.
Vol	DELAY1Y	-0.52	-1.66	Hou and Moskowitz (2005)	DELAY1Y equals $1 - \left(R^2_{\text{base}}/R^2_{\text{full}}\right)$, where R^2_{base} is estimated from the model $r_i - r_f = \alpha_i + \beta^i \left(Watch^{Dow} - r_f\right) + \epsilon_i$ and R^2_{full} from the same model but including four months of lagged $Watch^{Dow}$ exc. returns. The model uses past 12 months exc. returns and req. at least 10 obs.
Vol	DELAY2Y	-0.32	-0.99	Hou and Moskowitz (2005)	DELAY1Y but using monthly returns over the past 24 months req. at least 20 obs.
Vol	DELAY3Y	-0.27	-0.78	Hou and Moskowitz (2005)	DELAY1Y but using monthly returns over the past 36 months req. at least 24 obs.
Vol	COSKEW1Y	0.13	0.60	Harvey and Siddique (2000)	Coskewness, measured as the slope coefficient β^i_1 in $r_i - r_f = \alpha_i + \beta^i_0 \left(Watch^{Dow} - r_f\right) + \beta^i_1 \left(Watch^{Dow} - r_f\right)^2 + \epsilon_i$. The model uses past 12 months exc. returns and req. at least 10 obs.
Vol	COSKEW2Y	0.39	1.48	Harvey and Siddique (2000)	COSKEW1Y but using monthly returns over the past 24 months req. at least 20 obs.
Vol	COSKEW3Y	0.26	1.37	Harvey and Siddique (2000)	COSKEW1Y but using monthly returns over the past 36 months req. at least 24 obs.
Vol	SEASON	-0.08	-0.28	Heston and Sadka (2008)	Return seasonality, measured as the average of lag 12- and lag 24-month returns in excess of the returns of a price-weighted index of the according watch brand.

[a] r 1,0 (reversal) generates a sign. return of -1.52% (t-stat: -3.84) until Feb. 2020 and a sign. positive return of 2.94% (t-stat: 2.92) afterwards.
[b] Based on our extensive analysis on momentum (r 4,1) in Section 6.4.3 and Section 6.4.4, we suggest this strategy to generate significant return differences in the cross-section of luxury watch returns.

6.3.1 Size characteristics

For our first two size characteristics shown in Table 6.4, the differences in the average returns of the highest to the bottom quintile portfolios are an insignificant 0.11% for the monthly price (PRC) and 0.14% for the highest price of the month (MAXDPRC). Interestingly, the mean excess return decreases from the bottom to the fourth quintile and sharply increases in the top quintile. Given the fact that the price-weighted watch market return $Watch^{Dow}$ has an average return of 0.64% (0.60% in excess of the risk-free rate), it is not the extreme portfolio returns that are striking, but the comparable low returns in the mid portfolios (3) and (4).

Table 6.4: Luxury watch size-related strategy returns.

This table reports time-series averages of monthly price-weighted quintile portfolio returns in excess of the risk-free rate. Portfolios are updated each month based on PRC (log price in the portfolio formation month), MAXDPRC (maximum price in the portfolio formation month), and AGE (number of months listed on Chrono24.com since June 2009). The last two columns report the average 5-1 hedge return for two distinct subperiods. The sample period is July 2010 to March 2022. Newey and West (1987) robust t-statistics are given in parenthesis and */**/*** indicate significance at the 10%/5%/1% level.

	Quintiles							
	1 (Low)	2	3	4	5 (High)	5-1	5-1 (07/10 - 02/20)	5-1 (03/20 - 03/22)
PRC	0.61^{*}	0.45^{*}	0.29	0.26	0.72	0.11	−0.33	2.19
	(1.88)	(1.70)	(0.97)	(0.84)	(1.61)	(0.29)	(−1.45)	(1.30)
MAXDPRC	0.56^{*}	0.45^{*}	0.38	0.25	0.70	0.14	−0.32	2.23
	(1.74)	(1.69)	(1.29)	(0.78)	(1.56)	(0.34)	(−1.49)	(1.33)
AGE	0.28	0.87^{**}	0.62	0.76^{*}	0.62^{*}	0.34^{**}	0.48^{***}	−0.30
	(0.74)	(2.08)	(1.43)	(1.88)	(1.82)	(2.27)	(3.43)	(−0.78)

Prices for the cheapest luxury watches increased roughly 0.59% per month, which is not significantly different from the 0.71% for the most expensive ones, while prices for average watches only increased by half of that. A possible explanation could be Porter's incompatibility hypothesis and the phenomenon of being "stuck in the middle" (see Porter (1980)): The watch market is characterized by strongly differentiated products (in terms of prices) while at the same time a wide range of *relatively* low-priced products is available (see Kaschny et al. (2015) and Panel B of Table 6.1). In the figurative sense, luxury watches that neither exhibit a distinct price differentiation (high exclusivity as indicated by very high prices) nor a distinct cost leadership (affordable luxury watches for a wider target group) should be less profitable. Using portfolio sorts raises the question of appropriate breakpoints that could possibly bias our results. Fama and French (1992) use only NYSE stocks for size breakpoints to ensure that they are not dominated by relatively small firms listed on other exchanges. Similarly, our highly skewed watch prices result in quintile portfolio (PRC) sorts that comprise on average 51% of all Patek Philippe watches in the top quintile. In unreported robustness tests we calculate quintile breakpoints using only watches from Patek Philippe and/or Rolex, and both together augmented with Audemars Piguet. We observe insignificant 5-1 returns among all specifications and in all subperiods between 0.22% and 0.30%.

The observed insignificant size strategy based on monthly prices could also be attributed to our specific sample of 20 luxury watch brands. Because Chrono24.com displays price charts only for quite highly liquid traded watches having a high number of available offers, some very exclusive watches are not included in our sample. The brand Richard Mille (among others, e.g., Jacob & Co.) for example is known for highly limited watches with official retail prices for most of them above $500,000. The most expensive offer as of 03/21/2023 is a model RM 056 with an ask price of $5.6 million. 292 of the total 1,108 offers for Richard Mille watches refer to a RM 011 with prices between $200,000 and $985,000 (depends on the specific subvariants).[25] In fact, these watches are highly illiquid and are not included in our sample for that reason, which may influence our findings for the size strategy.

Sorting on the number of months listed on Chrono24.com (AGE), we observe significant differences in the average returns of the top and low quintile portfolio of 0.34%. Until February 2020, the difference was a highly significant 0.48% and plunged to an insignificant -0.30% afterwards. Without stepping into the dispute about the relevance of the equity size factor (e.g., van Dijk (2011), Asness et al. (2018), Hou and Dijk (2019)), we regard this finding at the very least as surprising because the size effect seems to be inverted. Unlike companies that decide on their own when to go public, it is Chrono24.com who decides what watches to track with historic price charts and when to start. This may distort the interpretation of watches with a long price-history as "big", established, watches. A plausible mechanism would be that Chrono24.com starts to track prices when a watch model becomes increasingly popular and recognized in public. We analyze this hypothesis using Google searches of individual watch models as a direct proxy for popularity and attention. We find a cross-sectional correlation between the deviation of Google searches compared with the average of those in the preceding three month and size (AGE) of 0.00. Similarly, the mean change of Google searches over past four months for watches that enter our sample is statistically not distinguishable from zero. Both results reject our supposition that Chrono24.com includes watches based on popularity, so we consider the

[25]At that day, Chrono24.com has a total of 530,332 watch offers, so Richard Mille accounts for only 0.21% of them.

interpretation of size (AGE) to be consistent with its counterpart in equities.

6.3.2 Value characteristics

Value is a more difficult characteristic to match to the market for luxury watches. Naturally, the counterpart to the book value of equity would be official list prices, which are generally not available for the history of our broad sample. Fama and French (1996) find that returns of portfolios generated by sorting on the negative of past 13-to-60-month returns and the book-to-market ratio are highly correlated. This view is supported by Gerakos and Linnainmaa (2018) who find that the value premium is specific to firms that become growth or value because they change in size - which is reflected in past returns - and not in book value. For that reason, we consider the negative of past 13-to-60-month returns as one of our measures for value (VAL_FF) in luxury watches.

Similarly, Asness et al. (2013) uses the log of the spot price, resp. spot exchange rate, five years ago divided by the most recent one, to define value in commodities and currencies. To be precise, they use the average of prices from 4.5 to 5.5 years ago as the price five years ago. We use this method as our second value variable (VAL_AMP).

Table 6.5: Luxury watch value-related strategy returns.

This table reports time-series averages of monthly price-weighted quintile portfolio returns in excess of the risk-free rate. Portfolios are updated each month based on the log of the average price from 4.5 to 5.5 years ago, divided by the most recent price (VAL_AMP, see Asness et al. (2013)), or the negative of the past 13-to-60-month return (VAL_FF, see Fama and French (1996)). The last two columns report the average 5-1 hedge return for two distinct subperiods. The sample period is July 2014 to March 2022 for VAL_FF and January 2015 to March 2022 for VAL_AMP. Newey and West (1987) robust t-statistics are given in parenthesis and */**/*** indicate significance at the 10%/5%/1% level.

	Quintiles							
	1 (Low)	2	3	4	5 (High)	5-1	5-1 (07/10 - 02/20)	5-1 (03/20 - 03/22)
VAL_AMP	1.58**	0.42	0.25	0.25	0.91	−0.67	−0.17	−1.90
	(2.06)	(1.18)	(0.76)	(0.69)	(1.53)	(−1.17)	(−0.36)	(−1.47)
VAL_FF	1.22	0.51	0.22	0.21	0.50	−0.72	−0.16	−2.25*
	(1.56)	(1.48)	(0.62)	(0.61)	(0.85)	(−1.42)	(−0.45)	(−1.89)

Table 6.5 shows the results for sorting luxury watches into portfolios based on their value characteristics. None of the zero-investment long-short strategies based on value generate statistically significant returns. We notice that portfolio returns decrease from the bottom to the fourth quintile portfolio and sharply increase in the top quintile. The only significant average return of 1.58% is generated by luxury watches in the lowest quintile portfolio according to VAL_AMP. The overall pattern among quintiles contrasts

with equities and signals that "growth" watches tend to have higher returns than "value" watches. Another interesting aspect is that returns in both extreme quintile portfolios are remarkably higher for VAL_AMP than for VAL_FF, although both have quite similar definitions. The crucial difference is that VAL_AMP includes returns over the past 12-months while VAL_FF only considers past 13-to-60-month returns. Luxury watches that tend to perform poor over the previous year tend to have a higher chance to be comprised in the top quintile portfolio in VAL_AMP based on the inverse long-term performance sorting. The fact that these watches generate returns of 0.91% instead of only 0.50% when omitting the performance of the last year as in VAL_FF indicates an inverse momentum effect, examined in more detail in the next section.

6.3.3 Momentum characteristics

In this section, we analyze the performance of the zero-investment long-short strategies based on past one-, two-, three-, four-, one-to-four-, eight-, one-to-twelve-, and 16-month returns. Each month, we sort individual luxury watches into quintile portfolios based on the value of a certain momentum characteristic and all strategies are rebalanced monthly. We find that two-, three-, four-, and 16-month momentum strategies result in significant (at least at the 5% level) negative long-short strategy returns (see Table 6.6). Up to past four-month momentum, the average mean excess return decreases in the portfolio quintiles apart from higher returns in the fourth quintile. For longer-term momentum strategies, we notice a U-shaped pattern among quintile portfolios.

The most striking result to emerge from the data is that all momentum strategies generate highly significant and economically huge positive returns in the *lowest* quintile portfolio. Investing in the lowest momentum quintile portfolio according to past two-, three-, four-, eight-, and 16-month returns not only significantly outperforms the highest quintile portfolio, but also our watch market portfolio $Watch^{Dow}$ by 0.75% per month on average. In other words, these low-momentum portfolio returns equal roughly twice the return of the watch market return.[26]

The relation between an asset's return and its recent relative performance is among the most studied capital market phenomena. To the best of our knowledge, all previous studies on momentum in equities (see Jegadeesh and Titman (2001)) for an extended discussion

[26] Momentum variables that include the most recent month are influenced by the reversal effect per definition. We analyze momentum measures skipping the return of the most recent month in more detail in Section 6.4.3 and Section 6.4.4. Our main conclusion of an inverse momentum effect is, however, not affected by this decision.

Table 6.6: Luxury watch momentum-related strategy returns.

This table reports time-series averages of monthly price-weighted quintile portfolio returns in excess of the risk-free rate. Portfolios are updated each month based on r a, b, which denotes the cumulative past b-to-a-month returns. The last two columns report the average 5-1 hedge return for two distinct subperiods. The sample period is July 2010 to March 2022 except for r 16,0 which begins in November 2010 because of limited data availability. Newey and West (1987) robust t-statistics are given in parenthesis and */**/*** indicate significance at the 10%/5%/1% level.

	Quintiles							
	1 (Low)	2	3	4	5 (High)	5-1	5-1 (07/10 - 02/20)	5-1 (03/20 - 03/22)
r 1,0	0.87^{***}	0.76^{**}	0.40	0.49	0.14	−0.73	-1.52^{***}	2.94^{***}
	(2.62)	(2.12)	(1.29)	(1.35)	(0.23)	(−1.41)	(−3.84)	(2.92)
r 2,0	1.38^{***}	0.60^{*}	0.36	0.50	0.04	-1.34^{***}	-1.98^{***}	1.64
	(3.32)	(1.88)	(1.15)	(1.38)	(0.07)	(−2.68)	(−5.02)	(1.43)
r 3,0	1.26^{***}	0.84^{**}	0.23	0.56	−0.04	-1.30^{***}	-1.87^{***}	1.33
	(2.62)	(2.44)	(0.72)	(1.61)	(−0.07)	(−2.61)	(−4.76)	(0.98)
r 4,0	1.35^{***}	0.61^{**}	0.24	0.50	0.02	-1.33^{**}	-1.99^{***}	1.73
	(2.92)	(1.97)	(0.77)	(1.41)	(0.03)	(−2.50)	(−4.52)	(1.51)
r 4,1	0.84^{**}	0.54^{*}	0.13	0.66	0.63	−0.21	-0.55^{*}	1.39
	(1.99)	(1.75)	(0.40)	(1.50)	(1.26)	(−0.60)	(−1.89)	(1.45)
r 8,0	1.37^{***}	0.78^{**}	0.15	0.49	0.30	-1.07^{*}	-1.70^{***}	1.85^{*}
	(2.95)	(2.06)	(0.46)	(1.40)	(0.51)	(−1.91)	(−3.35)	(1.71)
r 12,1	0.83^{**}	0.52	0.46	0.51	0.82	−0.01	−0.43	1.96^{*}
	(2.38)	(1.48)	(1.34)	(1.62)	(1.51)	(−0.03)	(−1.35)	(1.71)
r 16,0	1.26^{***}	0.25	0.20	0.24	0.37	-0.89^{**}	-1.41^{***}	1.45
	(3.22)	(0.82)	(0.70)	(0.66)	(0.63)	(−1.98)	(−3.57)	(1.42)

of the relevant literature), currencies (Kho (1996)), commodities (Erb and Harvey (2006) and Miffre and Rallis (2007)), international equity indices (Bhojraj and Swaminathan (2006)), residential real estate (Beracha and Skiba (2011)), credit default swaps (Lee et al. (2021)), and cryptocurrencies (Liu et al. (2022)) find significant positive future returns of past *winners*, i.e., high momentum characteristics. This is documented even for the case of examining momentum strategies not isolated for a single asset class, but jointly across diverse markets, as shown in Asness et al. (2013). A well-known exception with at least insignificant momentum strategies is Japanese equities (see Rouwenhorst (1998) and Griffin et al. (2003)), but Asness (2011) argues that momentum strategies in Japan are as successful as in other regions when combined with value-investing principles. In context of these studies, we document the first significant negative momentum effect in investible assets.

Our analysis in Section 6.2.2 reveals that the beginning of 2020 was a remarkable start of huge price increases for luxury watches. Looking at the according subperiods shows that all momentum strategies had even lower 5-1 portfolio returns until February 2020 compared to the full sample period. For the last two years in our sample, all momentum strategies generate insignificant return differences among quintile portfolios. To our surprise, reversal (r 1,0) is a highly significant positive 2.94% from March 2020 to March 2022 which means that past-month top performing watches tend to increase in prices again in the following month.

6.3.4 Volatility characteristics

We analyze the performance of the volatility-related return predictors described in Table 6.3 in this section. Using past 12-, 24-, or 36-months excess returns for estimating beta reveals that all high-low strategies generate insignificant difference returns over the entire sample period (see Table 6.7). Until February 2020, the price-weighted portfolio of low-beta watches (2Y) outperforms the highest quintile portfolio by a significant 0.38%. In line with previous studies on equities, low-beta watches tend to generate quite high and well measured returns, but there is no evidence of neither a positive, nor a negative, relation between beta and future watch-returns.[27]

Our results for idiosyncratic volatility (Ang et al. (2006)), delay (Hou and Moskowitz (2005)), and coskewness (Harvey and Siddique (2000)) are easily summarized. None of the according high-low strategies generates significant difference returns at the 5% level over the entire sample period. The overall return pattern of relatively low returns in the highest IVOL quintile portfolio is in line with results for equities. In contrast to the findings of Hou and Moskowitz (2005) for stocks, high delay watches generate the lowest returns among each quintile portfolios. We even document a highly significant difference return between the highest and lowest delay quintile portfolio of -2.22% since March 2020 when using past 24 months of returns for estimating delay. A possible explanation could be that the watch market portfolio started to generate very high returns at the beginning of 2020 and afterwards. Watches that drive this performance mechanically generate higher returns than high delay watches, where delay is referred to the watch market portfolio (and its low performance until the end of 2019). Results for coskewness reveals that systematic skewness is not related with future watch returns.

We define MAX to be the ratio of the maximum price in the portfolio formation month and the minimum price in the previous month. Our monthly prices are actually averages of the according daily prices for that month as described in Section 6.2.1. For that reason and in lack of daily data, we suggest that this definition best suits the common definition

[27] See e.g., Frazzini and Pedersen (2014), Novy-Marx and Velikov (2022), and the literature therein.

Table 6.7: Luxury watch volatility-related strategy returns.

This table reports time-series averages of monthly price-weighted quintile portfolio returns in excess of the risk-free rate. Portfolios are updated each month based on beta and squared beta (Fama and MacBeth (1973)), idiosyncracic volatility (Ang et al. (2006)), delay (Hou and Moskowitz (2005)), and coskewness (Harvey and Siddique (2000)), all estimated over past 12 (1Y), 24 (2Y), or 36 (3Y) months requiring at least 10, 20, resp., 24 observations. The last two rows show results for portfolios sorted on the maximum return in the portfolio formation month (MAX, Bali et al. (2011)) or on past two-year return seasonality (SEASON, Heston and Sadka (2008)). The last two columns report the average 5-1 hedge return for two distinct subperiods. The sample period starts in July 2010 for 1Y variables and MAX, July 2011 for 2Y variables and SEASON, and July 2012 for 3Y variables. DELAY begins in November 2013 because of data availability for lagged variables. The sample period for all variables ends in March 2022. Newey and West (1987) robust t-statistics are given in parenthesis and */**/*** indicate significance at the 10%/5%/1% level.

	Quintiles							
	1 (Low)	2	3	4	5 (High)	5-1	5-1 (07/10 - 02/20)	5-1 (03/20 - 03/22)
BETA1Y	0.68**	0.70*	0.53	0.70*	0.62	−0.05	−0.39	1.49
	(2.08)	(1.74)	(1.37)	(1.96)	(1.27)	(−0.18)	(−1.62)	(1.50)
BETA2Y	0.39	0.41	0.12	0.51	0.45	0.06	−0.38**	1.88
	(1.17)	(1.21)	(0.36)	(1.30)	(0.84)	(0.17)	(−2.20)	(1.24)
BETA3Y	0.55*	0.33	0.67	0.58	0.82	0.27	−0.22	2.07
	(1.75)	(1.00)	(1.31)	(1.37)	(1.36)	(0.65)	(−1.13)	(1.37)
$BETA1Y^2$	0.54*	0.77	0.53	0.79*	0.59	0.05	−0.39	2.11
	(1.78)	(1.64)	(1.40)	(2.06)	(1.20)	(0.16)	(−1.61)	(1.72)
$BETA2Y^2$	0.37	0.43	0.19	0.47	0.44	0.07	−0.41**	2.08
	(1.15)	(1.27)	(0.52)	(1.22)	(0.83)	(0.18)	(−2.34)	(1.25)
$BETA3Y^2$	0.59*	0.46	0.41	0.52	0.81	0.23	−0.25	1.99
	(1.83)	(1.16)	(1.23)	(1.25)	(1.35)	(0.57)	(−1.23)	(1.36)
IVOL1Y	0.73*	0.93**	0.85**	0.58*	0.26	−0.48*	−0.45	−0.62
	(1.77)	(2.17)	(2.10)	(1.65)	(0.64)	(−1.70)	(−1.42)	(−1.01)
IVOL2Y	0.45	0.79*	0.47	0.35	0.26	−0.18	0.10	−1.36**
	(0.96)	(1.75)	(1.25)	(0.94)	(0.61)	(−0.70)	(0.41)	(−2.12)
IVOL3Y	0.42	0.76*	0.92*	0.57	0.41	−0.01	0.22	−0.87
	(1.01)	(1.72)	(1.71)	(1.27)	(0.90)	(−0.03)	(0.77)	(−1.38)
DELAY1Y	0.72	0.69	0.83*	0.36	0.20	−0.52*	−0.22	−1.86**
	(1.53)	(1.64)	(1.82)	(1.05)	(0.59)	(−1.66)	(−0.74)	(−2.37)
DELAY2Y	0.59	0.68	0.37	0.38	0.27	−0.32	0.15	−2.22***
	(1.24)	(1.44)	(0.93)	(1.13)	(0.73)	(−0.99)	(0.67)	(−2.58)
DELAY3Y	0.58	0.76	0.82	0.42	0.42	−0.16	0.20	−1.41*
	(1.12)	(1.60)	(1.51)	(1.32)	(0.90)	(−0.49)	(0.66)	(−1.82)
COSKEW1Y	0.56	0.53	1.01*	0.66**	0.69*	0.13	−0.05	0.98*
	(1.43)	(1.53)	(1.85)	(2.02)	(1.73)	(0.60)	(−0.25)	(1.72)
COSKEW2Y	0.21	0.44	0.47	0.55	0.60	0.29	0.42*	0.29
	(0.53)	(1.12)	(1.01)	(1.47)	(1.38)	(1.37)	(1.78)	(0.30)
COSKEW3Y	0.43	0.69	0.34	0.78**	0.71	0.26	0.30	0.26
	(0.97)	(1.34)	(0.87)	(2.16)	(1.50)	(1.37)	(1.28)	(0.54)
MAX	0.48	0.71**	0.86**	0.97*	−0.17	−0.65**	−0.90***	0.52
	(1.30)	(1.99)	(2.21)	(1.88)	(−0.48)	(−2.30)	(−3.27)	(0.87)
SEASON	0.45	0.61	0.19	0.70	0.38	−0.08	−0.40*	1.29**
	(1.31)	(1.42)	(0.46)	(1.46)	(0.86)	(−0.28)	(−1.71)	(2.14)

of MAX for equities as described in Bali et al. (2011). A zero-investment strategy that longs high-MAX watches and shorts low-MAX watches generates a highly significant -0.65% per month. This result is, however, robust against other possible definitions. If we choose MAX to be the ratio of maximum and mean prices of the portfolio formation month, the strategy results in a significant difference return of -0.59% (t-statistics: -2.43). The latter definition only uses price information from the most recent month and therefore provides additional evidence that our results for MAX are not driven by the high returns of past-month losers (0.87%) as documented in the previous section. If we look at the returns across quintiles, we see average returns increase from 0.48% to 0.97%, but going from the fourth to the fifth quintile, average returns drop significantly from 0.97% to -0.17%. From July 2010 to February 2020, average returns across quintile 1-4 are within the range from 0.23% to 0.44% and then sharply decline to a significant -0.67% for the highest quintile. The observed return pattern for portfolios sorted on MAX is very similar to equities. An according MIN strategy (untabulated) in the luxury watch market generates only an insignificant difference return of -0.35%, which is also in line with findings on equity markets.

Heston and Sadka (2008) find that stocks tend to have relatively high (or low) returns every year in the same calendar month. We define return seasonality (SEASON) as the average of lag 12- and lag 24-month returns in excess of the returns of a price-weighted index for the according watch brand.[28] We detect only a weak significant average high-low SEASON difference return of -0.40% until February 2020 and a significantly 1.29% afterwards.[29] On average, the performance of a zero-investment long-short strategies based on SEASON is an insignificant -0.08%. We conclude that luxury watch returns are not predictable by long-run historical returns.

6.4 Additional results

In this section, we provide additional results for successful watch return predictors found in previous sections. First, all predictors were analyzed independently from each other, so we additionally test their joint significance. Second, we discuss potential explanations for our return predictors that have been proposed in the asset pricing literature for equities.

[28] Our findings are robust if we additionally include lag 36-month returns or calculate SEASON in excess of an equally weighted brand index.

[29] Results are similar if we use an equal-weighted index return as benchmark as in Heston and Sadka (2008). Including an additional lag 36-month return as a further robustness test also generates similar results, however, the high-low SEASON difference return from March 2020 to March 2022 becomes insignificant.

6.4.1 Joint significance of successful predictors

We find significant size, momentum, and MAX effects in the returns of luxury watches. Our zero-investment strategies are, however, not independent from each other, so we apply the k-FWER method of Lehmann and Romano (2005) to test their joint significance at the 10% and 5% levels. We find ten significant (at the 5% level) predictors between July 2010 and February 2020 and five between March 2020 and March 2022. Because we initially considered 30 characteristics, resp. hypothesis, we cut the 10% p-value threshold to $10\% \cdot 10/30 = 3.33\%$ for the first subperiod and to $5\% \cdot 10/30 = 1.67\%$ for the second subperiod. Similarly, we adjust the 5% p-value thresholds. All momentum strategies jointly remain significant at the 5% level under the k-FWER threshold. The same holds for size (AGE) and MAX, so the only strategies that are not significant at the 5% level are BETA and BETA squared, both using past 24 months as estimation period. In the second subperiod, only reversal (r 1,0) remains significant at the 5% level. In addition, a F-test with the null hypothesis that mean excess returns of all 30 strategies (successful or not) are jointly zero is rejected at the 1% level. Both the k-FWER method and the joint F-test provide evidence, that our results are unlikely to generate by chance.

6.4.2 Size effect and cost of arbitrage

To explore potential economic mechanisms for the size effect in luxury watch returns, we first explore the characteristics of size (PRC), i.e., absolute price in U.S.$, sorted quintile portfolios. The first quintile portfolio contains watches with an average price of $2,598. Watches comprised in this portfolio have an average idiosyncratic volatility (IVOL1Y) of 20.52% and an average coskewness (1Y) of -0.15, whereas watches in the highest quintile portfolio have an average idiosyncratic volatility of 14.61% and an average coskewness of 0.54. The average standard deviation of past 12-months returns in the lowest (highest) quintile is 7.28% (5.94%). These results are consistent with the common view of size in equities, that the strategy reflects risk in the comovement among stocks proxied for companies, that are smaller, more volatile, and higher exposed to idiosyncratic risks. Without actual data on transaction volumes, however, we are not able to directly discuss if size actually proxies for a potential illiquidity premium. If so, the size premium should be more pronounced among watches with higher arbitrage costs (see Shleifer and Vishny (1997) and Pontiff (2006)). In line with Stambaugh et al. (2015), we use a composite index to proxy for the cost of arbitrage. We sort watches in increasing order based on their idiosyncratic volatility and past 12-months return volatility into quintile portfolios. Similarly, we sort watches in decreasing order based on size (PRC) and size (AGE), since lower values of these variables indicate higher arbitrage costs. Each watch is given the corresponding score of its quintile rank for all four variables. The cost of arbitrage index on the watch level is the sum of the four scores and higher values indicate higher costs of arbitrage. Finally, each month we first sort watches into one of three cost of arbitrage portfolios (using 30th and 70th percentile breakpoints) and within each of these portfolios, we then sort watches based on their price. The long-short strategy for the tercile portfolio with high arbitrage costs generates a difference return of -1.04% (t-statistics: -2.49). The long-short strategy for the lowest cost of arbitrage tercile portfolio generates a statistically insignificant return premium of 0.49%. These results are consistent with the view that size proxies for an illiquidity effect, but do not imply a definitive answer.

6.4.3 Reversal, momentum, and MAX effects after controlling for size

Table 6.8 shows results for zero-investment strategies of successful predictors from the previous section after controlling for different price levels.[30] Using tercile instead of quintile portfolios ensures that the according sub-portfolios remain diversified and provides further robustness of our results. Breakpoints for size (PRC) are the according 30th and 70th percentile values.

In Panel A, we notice that most of the size (AGE) premium stems from the high returns of 0.67% for less expensive watches that are listed for a longer time on Chrono24.com. Watches with high MAX-characteristics (Bali et al. (2011)) generate returns that are not distinguishable from zero among all price (Panel B) and size (AGE) (Panel C) levels. Most of gains from the according zero-investment strategy stems from the relatively high returns for low MAX watches.

Past month losers (r 1,0) perform relatively well in the subsequent month, whereas small winner generate significant negative returns of -0.87%. Looking at short-term momentum measured as the cumulative return over past four-months (r 4,0), we surprisingly notice that watches in the lowest momentum tercile portfolios generate high returns between 1.07% and 1.94% among all size levels. Of course, this momentum measure includes the most recent month is driven by the strong reversal effect. To carefully separate momentum from reversal, we additionally show results for r 4,1 (past four-to-one month) in Panel F. Without any bias induced by one-month reversal, we observe that zero-investment

[30]Size (PRC) does not generate significant zero-investment strategy returns as shown in Table 6.4, whereas size (AGE) does. We choose size (PRC) as our control variable because of its straightforward interpretation of being price levels of luxury watches and being very easily observable. Second, size (AGE) depends purely on the incomprehensible decision of Chrono24.com whether at all, and if so, when to start tracking prices of certain watches.

Table 6.8: Luxury watch portfolio strategy returns controlling for size effects.

This table presents results of bivariate independent-sort portfolio analyses for successful predictors of the cross-section of luxury watch returns after controlling for size. Each month, all luxury watches are independently sorted into three groups (small, middle, and big) based on an ascending sort on size (PRC) (Panel C: size (AGE)), i.e., the price of a watch in U.S.-$ at the portfolio formation month, and size (AGE), reversal, momentum, and MAX. Size (PRC) is the price in U.S.-$ at the portfolio formation month and size (AGE) the number of months listed on Chrono24.com. Reversal (r 1,0) is past one-month return and momentum (r 4,1) is the cumulative past four-month return, skipping the most recent month. MAX denotes the ratio of the maximum price in t_0 and the minimum price in t_{-1}. Breakpoints for all variables are the according 30th and 70th percentile values. The intersections of the 3 × 3 independent sorts produces nine portfolios, and we calculate their price-weighted returns in excess of the risk-free rate (one-month Treasury bill) for the subsequent month. The sample period is July 2010 to March 2022. We provide Newey and West (1987) robust t-statistics and */**/*** indicate significance at the 10%/5%/1% level.

	Avg. excess returns				t-statistics			
	1 (Low)	2	3 (High)	3-1	1 (Low)	2	3 (High)	3-1
Panel A: Size (PRC) - Size (AGE) portfolios								
Small	0.14	0.48	0.67**	0.53**	0.39	1.60	2.26	2.02
Middle	0.30	0.30	0.41	0.11	1.02	0.96	1.35	0.48
Big	0.44	0.80*	0.80*	0.36*	1.07	1.77	1.90	2.01
Panel B: Size (PRC) - MAX portfolios								
Small	0.50	0.52	0.38	−0.12	1.47	1.41	1.25	−0.46
Middle	0.41	0.47	0.01	−0.40*	1.26	1.58	0.04	−1.75
Big	0.69	0.92*	0.09	−0.60*	1.42	1.95	0.25	−1.81
Panel C: Size (AGE) - MAX portfolios								
Small	0.47	0.69*	0.01	−0.47*	1.22	1.68	0.02	−1.87
Middle	0.90**	0.94**	0.00	−0.90***	2.05	2.08	0.00	−2.89
Big	0.65*	0.81**	0.32	−0.33	1.69	2.13	0.98	−0.88
Panel D: Size (PRC) - Reversal (r 1,0) portfolios								
Small	1.31***	0.65*	−0.87**	−2.17***	3.91	1.94	−2.21	−5.02
Middle	1.02***	0.30	−0.29	−1.32***	3.15	0.93	−1.03	−5.39
Big	0.92*	0.33	0.47	−0.45	1.80	1.04	0.78	−1.02
Panel E: Size (PRC) - Momentum (r 4,0) portfolios								
Small	1.94***	0.38	−1.00***	−2.94***	5.00	1.18	−2.67	−6.30
Middle	1.18***	0.42	−0.59**	−1.77***	3.30	1.43	−1.99	−6.88
Big	1.07**	0.44	0.41	−0.65	2.02	1.33	0.72	−1.29
Panel F: Size (PRC) - Momentum (r 4,1) portfolios								
Small	1.40***	0.37	−0.17	−1.57***	3.85	1.19	−0.46	−5.09
Middle	0.56*	0.45	−0.02	−0.58***	1.72	1.51	−0.06	−2.99
Big	0.67	0.48	0.81	0.14	1.47	1.42	1.54	0.39

momentum strategies generate highly significant negative returns of -1.57% for less expensive (small) watches and -0.58% for the mid-size group. This is surprising, because our analysis of univariate quintile portfolios sorted on r 4,1 only generated insignificant return differences of -0.21%. This discrepancy is attributed to the economically strong but statistically insignificant returns of 0.81% among most expensive, high momentum watches. Using a longer estimation period for momentum, skipping the most recent month, reveals insignificant difference returns among both the small and big category. The only exception is a highly significant return of -0.90% in the middle size segment among r 12,1 sorted luxury watches. We conclude that momentum, unaffected by short-term reversal of past one-month returns, is highly prevalent in the short-term of four month and tends to be much weaker for longer estimation periods.

6.4.4 Momentum and investor attention

Explanations for investment strategies based on past returns often involve behavioral models and are based on the inability of investors to reflect new information instantaneously and correctly in asset prices. Two possible channels extensively discussed in the literature are delayed overreaction and initial underreaction. In the model proposed by Daniel et al. (1998), investors overreact to information, resp. react delayed, because of their overconfidence. Empirically testable implications are that returns positively covary in medium-term (Jegadeesh and Titman (1993)) and mean-revert in the long-run (De Bondt and Thaler (1985)). Another possible channel discussed in Peng and Xiong (2006) links initial overreaction to investor attention. What follows is that momentum should be more prevalent among assets that receive high attention. In the model of Barberis et al. (1998), momentum is based on initial underreaction. Investors do not fully incorporate new information into prices. The mechanism for initial underreaction in the model of Hong and Stein (1999) assumes that firm-specific information diffuses gradually across investors, who observe different private information at different points in time. Our results in Table 6.8 that momentum r (4,1) is more pronounced among small, i.e., less expensive, luxury watches are in line with underreaction-based explanations similar to the equity market (Hong et al. (2000)). In both cases of underreaction, we would not observe long-term reversal in momentum returns.

To begin with, Figure 6.4 displays the cumulative return in excess of the risk-free rate over the months $t + 1$ to $t + 36$ for portfolios that are sorted into quintile portfolios in month t on momentum r 4,1. We observe that holding the initial winner portfolio up to five months generates on average a cum. exc. return of -1.35%. Then, the initial

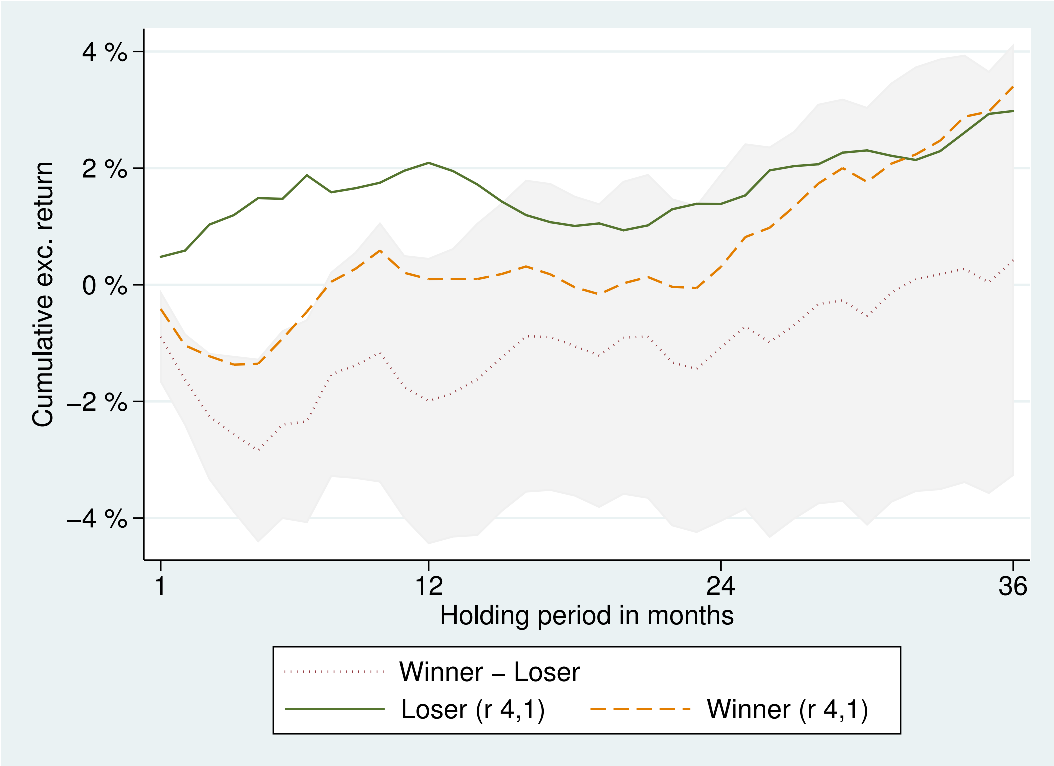

Fig. 6.4. This figure displays the cumulative return in excess of the risk-free rate over the months $t+1$ to $t+36$ for portfolios that are sorted into quintile portfolios in month t on past four-to-one month returns (momentum r 4,1). Winner (Loser) denotes the highest (lowest) quintile portfolio and returns are price weighted. The sample period is July 2010 to March 2022. The shaded area represents the confidence interval at the 99% level for the zero-investment strategy return that is long the highest quintile and short the lowest quintile momentum portfolio. We use White (1980) robust standard errors for the confidence interval.

underperformance reverts to a cum. exc. return of 0.21% up to a holding period of one year. The lowest quintile momentum portfolio (loser) generates a cum. exc. ret. of 2.10% over one year that only slightly reverts to a total of 1.39% in the subsequent year. The null hypothesis of Winner-Loser difference returns being zero is rejected for holding periods up to seven months at the 99% confidence level. Our results show that the initial negative returns of past winners mean revert, but interestingly, the cum. exc. returns of past losers remain above their initial response in $t+1$ for all subsequent holding periods.

If momentum is driven by initial overreaction, it should be more pronounced among high-attention watches. We test both, the cross-sectional and the time-series implications of this hypothesis. Following Liu et al. (2022), we use worldwide Google web searches to proxy for investor attention. Specifically, our measure for attention is the deviation of Google searches in a given month t_0 compared with the average of those in the preceding three months t_{-4} to t_{-3}. We standardize the Google search measure to have a mean of zero and a standard deviation of one. In contrast to previous studies, we measure Google searches on the more granular watch level instead of an aggregated level. For each watch in our sample, we use the combination of brand name and watch model as search term. We adjust this term in case that no search results could be obtained but carefully consider that the term reflects a unique watch model. As an example, the term *"Breitling Superocean Chronograph M2000"* does not return search results, whereas dropping the word *"Chronograph"* does. This, however, does not change the unique character of the specific watch model searched for.[31] Each month, we independently sort luxury watches into three groups based on our Google search measure and on momentum (r 4,1), using the 30th and 70th percentile of these characteristics as breakpoints. We calculate price-weighted portfolio returns in excess of the risk-free rate for the subsequent month.

In contrast to the hypothesis of Peng and Xiong (2006), our cross-sectional portfolio analysis reveals that the momentum (r 4,1) effect is more pronounced among luxury watches with low attention. For high-search group, the long-short strategy return is a statistically insignificant 0.46% per month. Past loser with low attention generate a significant exc. return of 1.13% and the Winner-Loser difference return is a significant -0.87%.[32] As discussed in Section 6.2.1, Rolex is considered to be the most prominent watch brand in our sample, and we suggest that their watches receive relatively high attention in public. Similarly, we observe an insignificant zero-investment strategy return of -0.20% (-0.23% from July 2010 to February 2020 and 0.73% from March 2020 to March 2022, all not significantly different from zero) when being long the top quintile of momentum (r 4,1) Rolex watches and shorting the according lowest quintile portfolio of Rolex watches. Further, the estimated slope coefficient in a pooled-regression of one-month ahead watch excess returns on Google searches is statistically not significantly different from zero (using robust standard errors clustered by both watch and month).

For the time-series analysis on momentum and attention, we aggregate the Google search measure by price-weighting the watch-specific measures. To examine if momentum is stronger at times of high investor attention, we regress momentum returns $r_t^{Win.-Los.}$

[31] Similarly, we often drop the numeric indication for the case size, e.g., *"Breitling Chronomat"* instead of the original term *"Breitling Chronomat 44"*, or the additional word *"GMT"*, referring to the functionality of quickly adapting new time zones. A total of ten watches is excluded for that analysis because of missing Google search results. These watches are Rolex "Athlete", A. Lange & Söhne "Kleine Lange 1" and "Grosse Lange 1", Girard Perregaux "WW.TC" and "Rattrapante Chronograph", Piaget "Rectangle A L Ancienne", Zenith "Pilot Cronometro Tipo Cp-2", and Chopard "Two O Ten".

[32] Results are similar if we use dependent portfolio sorts.

from quintile portfolio sorts on the one-month lagged aggregated Google search measure $google^{agg}_{t-1}$ and on the interaction between $google^{agg}_{t-1}$ and a dummy variable $TIME_t$ that is zero for all months until February 2020 and one for months March 2020 to March 2022 (Newey and West (1987) robust t-statistics are provided in parenthesis):

$$r^{Win.-Los.}_t = \underset{(-1.91)}{-0.55} + \underset{(1.87)}{2.02} \cdot TIME_t + \underset{(-0.05)}{-0.04} \cdot google^{agg}_{t-1} + \underset{(1.47)}{5.68} \cdot google^{agg}_{t-1} \cdot TIME_t + \epsilon_t. \quad (6.1)$$

The estimated regression coefficients imply that zero-investment momentum strategy returns are not driven by investor attention. Momentum returns are on average 2.02 percentage points higher after February 2020. A one standard deviation increase in $google^{agg}$ after February 2020 increases momentum returns by an economically large 5.68 percentage points per month, which is nevertheless statistically insignificant. As an additional robustness test, we reapply the regression in Eq. (6.1) using a directly aggregated version of $google^{agg}_{t-1}$, that is based on the search term *"luxury watches"* and observe similar results.

6.4.5 Luxury watch returns and investor sentiment

If returns of a zero-investment strategy are, at least partially, associated with mispricing, they should be affected by time-variation in sentiment (Baker and Wurgler (2006)). Generally, the primary form of mispricing is overvaluation because it is harder to correct due to short-selling restrictions (Stambaugh et al. (2012)). During high-sentiment periods, the optimistic projections on future prices tend to be overly optimistic, and assets are more likely to be overpriced. On the other hand, projections tend to be more realistic in low-sentiment periods, and assets are more likely to be priced correctly.

We use the monthly index presented in Baker and Wurgler (2006) to proxy for investor sentiment.[33] We determine the relation between luxury watch returns and sentiment effects with time-series regressions on one-month lagged levels of sentiment (SENT), a dummy variable (TIME) that takes the value of one for all months between March 2020 and March 2022 and zero otherwise, and an interaction term between SENT and TIME:

$$r_{i,t} - rf_t = \alpha_i + \beta_{i,1} TIME_t + \beta_{i,2} SENT_{t-1} + \beta_{i,3} SENT_{t-1} \times TIME_t + \epsilon_t. \quad (6.2)$$

We separately analyze the returns of the lowest (highest) quintile portfolio for our

[33] The index is constructed to proxy for sentiment in the U.S. stock market, but international investor sentiment is primarily driven by U.S. sentiment as documented in Baker et al. (2012).

successful return predictors size (AGE), MAX, reversal r (1,0), and momentum r (4,1), and results are shown in Table 6.9.

Table 6.9: Luxury watch returns and investor sentiment.

This table presents coefficient estimates for monthly time-series regressions of zero-investment strategy returns in excess of the risk-free rate on one-month lagged investor sentiment (SENT), a dummy variable (TIME) that takes the value of one for all months between March 2020 and March 2022 and zero otherwise, and an interaction term between SENT and TIME. We use the index presented in Baker and Wurgler (2006) to proxy for investor sentiment. Each month, all luxury watches are sorted into five quintile portfolios based on size (AGE), MAX, reversal (r 1,0), and momentum (r 4,1), and we calculate their price-weighted returns in excess of the risk-free rate for the subsequent month. See Table 6.3 for variable definitions. Short (Long) leg referes to the exc. returns of the lowest (highest) quintile portfolio. The sample period is July 2010 to March 2022. We provide Newey and West (1987) robust t-statistics in parenthesis and */**/*** indicate significance at the 10%/5%/1% level.

	Short Leg				Long Leg			
	AGE	MAX	r (1,0)	r (4,1)	AGE	MAX	r (1,0)	r (4,1)
Intercept	−1.15**	−0.89	−0.33	−0.50	−0.62	−0.86	−1.10	−0.82
	(−2.17)	(−1.60)	(−0.58)	(−0.72)	(−1.04)	(−1.24)	(−1.40)	(−1.05)
TIME	2.14***	0.92	0.49	0.52	1.60*	1.41	2.02*	1.28
	(2.94)	(1.14)	(0.51)	(0.44)	(1.77)	(1.41)	(1.68)	(1.16)
SENT	−4.05**	−4.82**	−4.49**	−4.38*	−3.82*	−0.82	−1.30	−3.40
	(−2.19)	(−2.35)	(−2.31)	(−1.77)	(−1.79)	(−0.35)	(−0.53)	(−1.35)
SENT × TIME	5.87***	6.68***	6.13***	7.07**	5.29**	2.70	5.50*	7.19***
	(2.98)	(2.90)	(2.91)	(2.46)	(2.36)	(1.08)	(1.99)	(2.60)

The estimated slope coefficients on both portfolio legs are negative, consistent with overall sentiment effects. The coefficients for the short legs are larger in absolute magnitude, indicating that the according watches are more affected by investor sentiment. The greater short-leg sensitivity is in line with arbitrage asymmetry discussed in Stambaugh and Yuan (2017) and indicates a systematic component of mispricing, because it leaves more uncorrected overpricing than uncorrected underpricing. If sentiment affects prices, then periods of high (low) sentiment are likely to be followed by especially low (high) returns on overpriced (underpriced) luxury watches. This is exactly what we observe here for all our successful predictors size (AGE), MAX, reversal r (1,0), and momentum r (4,1), and is in line with a mispricing interpretation, i.e., that these strategies reflect a mispricing commonality across luxury watches. Analyzing the aggregated watch market by using $Watch^{Dow}$ exc. returns as dependent variable, we observe:

$$Watch_t^{Dow} - rf_t = \underset{(-1.12)}{-0.67} + \underset{(1.49)}{1.34} \cdot TIME_t + \underset{(-1.66)}{-3.51} \cdot SENT_{t-1} + \underset{(2.58)}{5.90} \cdot SENT_{t-1} \times TIME_t + \epsilon_t. \tag{6.3}$$

One unanticipated finding is that returns of both the lowest and highest quintile portfolios are positively related with sentiment after February 2020, as the sum of coefficients $SENT$ and $SENT \times TIME$ significantly exceeds zero. Especially short-leg returns increase higher in magnitude than their long counterpart, which results in a decrease of their difference returns as documented in previous sections. Even the exc. returns of the luxury watch market portfolio increase on average by 2.39 percentage points per month after February 2020 for a one unit increase in sentiment. The level of the sentiment index strongly increased from -0.08 (February 2020) to a maximum of 2.28 in December 2021 and we see remarkably high returns for luxury watches during that period. In the context of the noise-trader model of De Long et al. (1990), there are two contrary channels affecting prices. First, investors generate upwards price pressure because of higher sentiment which results in lower expected returns afterwards. Second, their increased demand for assets amplifies overall market risk ("hold-more effect") resulting in higher expected returns. Our significantly positive coefficient estimates imply that the hold-more effect dominates in the short-run, thus sentiment affects watch prices in equilibrium. Surging increase in investor sentiment over time has a strong predictive power of a plunge in asset prices in subsequent periods (Pan (2020)). Taken together, the initial increase in prices is conceivably not related to fundamental signals resulting in common mispricing, complementing the previously detected arbitrage asymmetry. This view is further supported by the results of Baker and Stein (2004) that irrational investors add liquidity to the stock market, only when they are optimistic, which is reasonable for the second half-year 2020 as indicated by the increasing level of investor sentiment. This implies that high liquidity is a symptom of overvaluation. Lower returns are actually observed after times of high liquidity as documented in Amihud (2002). The amount of Rolex watch models - the most extensive in our sample - listed for sale on Chrono24.com increased by 4.8% in 2020 compared to the previous year (Andrioli (2020)). In their report on the state of the global watch industry, Chrono24 (2022) reports that the number of sales in the first eight months of 2022 was 19% higher than the previous year. The total sales volume increased by 42%. These findings support our view that the remarkable price increase since February 2020 is mostly driven by mispricing.[34]

[34] Another possible explanation for our findings could be related with wealth effects as in Dimson and Spaenjers (2011), i.e., changes in wealth of affluent individuals can be expected to drive the market return for watches. We object this view for two reasons: First, the correlation between $Watch^{Dow}$ and global, developed stock market returns sharply decline after February 2020 to a negligible 0.10. Second, the equity

Indeed, there is strong evidence that prices sharply decreased in times after our sample period until March 2022: The Financial Times article *"Is that the sound of a luxe watch bubble popping?"* (Elder (2023)) reports that average prices for Rolex watches on secondary markets declined by 15% between January 2022 and January 2023. Millenary Watches, a large retailer of pre-owned luxury watches, states, that *"Some luxury watches dropped as much as 15-30% of their value from the peak, which in hindsight can be said was right around January-February of 2022. [...] And the watches that have increased the most during this time seem to be the watches that have taken the biggest hit (surprise!)"* (see Millenary Watches (2023)). Bloomberg reports that the *Subdial50 Index* of most traded watches lost 33% within a year until February 2023 and Chrono24.com cuts 13% of its workforce (Hoffman (2023)).

6.4.6 Luxury watch returns and stock market factors

Asness et al. (2013) finds that zero-investment strategies covary across different asset classes which raises the question if our luxury watch strategies also covary with their counterpart in equity markets. We use factor returns for the Fama/French five-factor model with an additional momentum factor (FF6) as proposed in Fama and French (2015) and Fama and French (2018) from Kenneth French's website. Daniel et al. (2020a) (DHS) introduce a model that aims to capture covariances with common elements of mispricing. Their two factors, FIN and PEAD, related to mispricing of a persistent nature (FIN) and of a transient nature (PEAD). The mispricing model by Stambaugh and Yuan (2017) (SY) also introduces two factors, PERF and MGMT, based on two distinct clusters of eleven prominent anomalies. Hanauer (2020) constructs SY factors for global, developed non-U.S. markets. We obtain these monthly factor returns from the according authors' websites. Our sample period is July 2010 to December 2016 (2018) for SY (DHS), resp. March 2020 for global, non-U.S. SY and FF6, which is determined by availability of the factor returns.

$Watch^{Dow}$ strongly covaries with both U.S. and global equity market returns and has a significant alpha of 2.01% per month after February 2020. There is a highly significant loading of -0.46 with the global, non-U.S. SY size factor. Luxury watch size (AGE) and MAX have highly significant alphas between July 2010 and February 2020 and the latter weakly comoves with the SY size factor. Results for momentum r (4,1) are ambiguous

premium should be more reasonable in terms of risk aversion when consumption is based on luxury goods as documented in Aït-Sahalia et al. (2004). We observe a correlation between real $Watch^{Dow}$ returns, deflated using CPI, and the monthly log change in U.S. real watch consumption (NIPA data on PCE Watches) of only 0.0067 for our sample period. This results in an implausible high level of risk aversion.

at first glance. As expected, we observe insignificant alphas for models FF6 and global, non-U.S. SY, because the univariate sorted quintile portfolio strategy also generates insignificant difference returns of -0.21%. On the other hand, we detect a highly significant negative alpha of -1.16% (t-statistics: -4.02) for SY and -0.82% (t-statistics: -3.09) for DHS. Exposure towards PERF is a weakly significant 0.21. This could be related to our findings in Section 6.4.3, that momentum is not prevalent among most expensive watches, which dominate price-weighted portfolio returns.

Reversal r (1,0) has a highly significant average alpha of -1.71% across all factor models until February 2020 and 3.21% afterwards, very similar to the results for our portfolio analysis presented in Table 6.6. Reversal also tends to negatively comove with FIN, which captures predominantly longer-term mispricing and correction. This is in line with our mispricing interpretation because the high-low difference return of r (1,0) is entirely driven by the high returns of past losers. In other words, watches in the lowest reversal quintile portfolio generate high returns because they are considered to be undervalued in terms of FIN and the already long-term undervaluation gets corrected in the subsequent month.

Overall, the factor-model adjusted alphas of luxury watch investment strategies are quantitatively similar to our previous results, and we conclude that comovement with equity markets does not sufficiently explain them.

6.4.7 Transaction costs

Other studies on the performance of collectibles often ignore transaction costs and refer to the assertion proposed in Goetzmann (1993) that these items have some non-monetary dividend yield considering the aesthetic pleasure and the social status of the owner.[35] Thus, non-monetary returns and the excessive costs associated with related investments are often assumed to be balanced. From the perspective of the asset pricing literature, this assumption does not hold, and we address the concern of neglecting transaction costs in this section.

Following Novy-Marx and Velikov (2016), we evaluate an easy-to-implement, rule-based, long-only buy/hold strategy designed to mitigate transaction costs when trading the luxury watch portfolio strategies discussed in Section 6.3: In the meaning of "sS rules" (see Arrow et al. (1951) and Davis and Norman (1990)), traders buy into a luxury watch

[35] See also Baumol (1986) and Mandel (2009) who emphasize the role of non-pecuniary payoffs and related non-financial utilities for collectors.

when it enters the buy range, but do not sell the watch until it falls out of the hold range, which is larger than the buy range. We implement a conservative 80%/60% buy/hold rule which implies that a trader investing in the long leg of an anomaly, buys watches when they enter the top 80th percentile of a given watch characteristic, and holds these watches until they fall below the top 60th percentile. Similarly, if one invests in the short leg of an anomaly, the trader buys watches when they are within the low 20th percentile of a given watch characteristic, and holds these watches until they rise above the low 40th percentile.[36]

Taking transaction costs into account of our 80%/60% buy/hold sS trading rule for each anomaly is quite easy. Depending on the number of luxury watches listed, Chrono24.com charges a fixed fee billed monthly for listing luxury watches.[37] Listing up to 25 watches costs 199 Euro, followed by a fee of 369 Euro for listing 26-50 watches. Listing up to 350 watches (i.e., selling the entire watch sample) is charged with 1,299 Euro.[38] These offered rates are automatically up- and down-scaled, i.e., if the average number of active listings exceeds (or falls below) the current package by more than 10% for two consecutive months, Chrono24.com will automatically adjust the package size. Based on the number of sells for trading a given luxury watch market anomaly, we will consider this fee structure as appropriate transaction costs for listing the watch on the marketplace. To be rather conservative, we additionally impose a buying-related fee of 1% of the price of a luxury watch when it first enters the portfolio. While this value is somehow arbitrary, it should reflect non-transaction costs like insurance and storage (see Le Fur and Outreville (2019)).[39]

[36]Investing in the short leg of an anomaly does not imply short-selling, but rather investing in a certain side of a characteristic. Take e.g., the size-anomaly among stocks, where one has to invest in the portfolio comprising low market cap firms to exploit the strategies' returns.

[37]This is the case for professional dealers that have at least 50 watches to sell with an average value of $2,000 each.

[38]To be precise, fees are 199 Euro up to 25 watches, 369 Euro up to 50 watches, 499 Euro up to 75 watches, 629 Euro up to 100 watches, 829 Euro up to 150 watches, 1,069 Euro up to 250 watches, 1,299 Euro up to 350 watches, 1,549 Euro up to 500 watches, 1,899 Euro up to 700 watches, and 2,199 Euro up to 1,000 watches.

[39]As shown in Table 6.2, the average price of a luxury watch in our sample is $13,614, so the average fee imposed for buying is $136.14, which seems reasonable to cover the costs for safe deposit boxes. To put these numbers into perspective, an all-risk insurance for a watch worth $13,614 costs $5.43 per month or less than 0.5% of the watch price annually on Segurio, a popular insurance underwriter offering valuable items insurance.

Table 6.10: Luxury watch returns net of transaction costs.

This table reports the performance of successful predictors for the cross-section of luxury watches (see Table 6.3), net of transaction costs. Following Novy-Marx and Velikov (2016), we evaluate an easy-to-implement, rule-based, long-only 80%/60% buy/hold strategy. The trading rule implies that a trader investing in the long leg of an anomaly, buys watches when they enter the top 80th percentile of a given watch characteristic, and holds these watches until they fall below the top 60th percentile. Similarly, if one invests in the short leg of an anomaly, the trader buys watches when they are within the low 20th percentile of a given watch characteristic, and holds these watches until they rise above the low 40th percentile. Transaction costs for selling are reflected in accounting for tiered fees for watch listings on Chrono24.com (see the main text for details). We additionally impose transaction costs for buying a luxury watch amounting for 1% of its price in the month of buying the watch. Portfolios are updated each month and we calculate price-weighted returns over the subsequent month. The first column refers to the anomaly traded and the second column reports the leg usded for the trading strategy. Persistence is calculated as the average cross-sectional correlation between the anomaly characterisitc in month t and month $t + 12$. Volume is the average dollar amount invested to pursue the trading strategy. The remaining columns report the average turnover (separately for buys and sells, in % of the portfolio value), the strategies' average gross return, and net return for the sample period from July 2010 to March 2022, and subperiods. Newey and West (1987) robust t-statistics are given in parenthesis and */**/*** indicate significance at the 10%/5%/1% level.

	Port.	Persistence in %	Volume in $	Turnover (in %) buys	Turnover (in %) sells	Gross return	Net return (07/10 - 03/22)	Net return (07/10 - 02/20)	Net return (03/20 - 03/22)
AGE	High	100.00	379,658	1.39	0.00	0.66*	0.64*	0.30	2.23***
						(1.94)	(1.90)	(0.90)	(3.31)
r 1,0	Low	30.76	597,561	60.09	58.22	1.15***	0.50	0.16	2.09
						(2.68)	(1.19)	(0.48)	(1.26)
r 2,0	Low	29.56	630,336	41.89	39.99	1.21***	0.76**	0.51	1.88
						(3.27)	(2.08)	(1.51)	(1.57)
r 3,0	Low	29.70	632,895	33.10	31.19	1.18***	0.81*	0.52	2.16
						(2.76)	(1.92)	(1.34)	(1.49)
r 4,0	Low	28.79	634,398	28.38	26.35	1.31***	0.98**	0.76**	1.98
						(3.20)	(2.44)	(2.05)	(1.41)
r 4,1	Low	29.76	632,245	33.11	31.19	0.84**	0.47	0.15	1.96
						(2.12)	(1.20)	(0.42)	(1.48)
r 16,0	Low	38.09	575,496	14.33	11.93	0.98***	0.76**	0.41	2.32***
						(2.79)	(2.17)	(1.21)	(2.65)
MAX	Low	52.64	895,635	16.23	14.21	0.57*	0.38*	0.12	1.56***
						(1.82)	(1.67)	(0.69)	(3.84)

Table 6.10 provides insights into our 80%/60% buy/hold portfolio strategies, trading luxury watches based on successful cross-sectional predictors found in Section 6.3. Trading the short leg of momentum, i.e., buying past-loser watches, requires more than \$600,000 to replicate the price-weighted portfolio strategy, substantially more than trading the size anomaly (AGE). Unexpected, the portfolio turnover decreases with longer momentum formation periods. Each month, around 30% of the portfolio is bought and/or sold.[40]

Looking at anomaly portfolio returns net of transaction costs, we notice that they amount for a substantial portion of gross returns. Because of the higher portfolio turnover, the r 1,0 momentum strategy, generating a highly significant 1.15% p.m. gross returns, turns an insignificant 0.50% net returns. However, the economic magnitude and the statistical significance for longer formation period momentum strategy returns remains: The r 4,0 strategy generates a substantial 0.98% p.m., most of it until February 2020. The size strategy yields a weakly significant 0.64%, but a highly significantly 2.23% net return since March 2020. The low-MAX strategy also generates a highly significantly 1.56% since March 2020. In conclusion, transaction costs substantially decrease the returns of our anomaly portfolio returns, but most of them still generate economically large and statistically significant net returns.

6.5 Conclusion

Our broad sample of 27,289 watch-month observations from the world's largest peer-to-peer marketplace for luxury watches Chrono24.com opens new possibilities to test theories of cross-sectional asset pricing anomalies. We are the first to test 30 characteristics related with the categories size, value, momentum, and volatility in the cross-section of 345 distinct luxury watches from 20 brands for the period June 2010 to March 2022.

Similar to other asset classes (see, e.g., Asness et al. (2013)), we find that size (AGE, i.e., the number of months listed on Chrono24.com, reversal (past one-month return), short-term momentum (past four-to-one month return), and MAX (Bali et al. (2011)) generate significant difference returns among zero-investment quintile portfolio strategies. Both the k-FWER test method by Lehmann and Romano (2005) and an F-test for the joint significance provide evidence that our results are unlikely to generate by chance. Overall, our findings are more in favor of a mispricing related interpretation and that the

[40]The reason for the low turnover of the size strategy AGE is that Chrono24.com shows price charts (i.e., our data source) only for popular watch models. Once displayed, none of the charts discontinuous, so the portfolio composition remains and comprises watches whose price charts are shown for the longest possible time period.

strategies reflect a mispricing commonality across luxury watches, rather than risk-based explanations.

Our objective is to use standard empirical asset pricing tools to analyze the cross-section of luxury watch returns and to provide out-of-sample insights into the understanding of anomalies in financial markets. While this study reveals new evidence consistent with mispricing as at least a partial explanation for our studied watch-counterparts to prominent equity anomalies, we do not aim to find complete explanations for each of them. Instead, studying the early-stage market for luxury watches helps us to establish a set of empirical regularities which can be used as stylized facts to asses theoretical asset pricing models, and, even more important, helps us to understand the dynamics of new upcoming markets for other collectibles, too. With more than 700,000 daily visitors in 2025, the marketplace for luxury watches provided by Chrono24.com is nascent and is constantly undergoing major transformations. Still, studying the dynamics of these arising investment opportunities even in an early stage appears to be a promising endeavor, as the case of cryptocurrencies has already shown in recent years.[41]

[41] The first finance-related academic studies analyzing the market for cryptocurrencies (back then only Bitcoin) appeared around 2012, despite high market frictions and its underdeveloped state with lots of speculations and fraud present. The first Bitcoin was created just a few years earlier, in 2009.

Appendix for
"The Global Market for Luxury Watches and Asset Pricing"

Abstract

- Section A.I: End-of-month versus mean-of-month price data.
- Section A.II: Luxury watch returns compared with fine wine and art returns.

A.I End-of-month versus mean-of-month price data

Volatility bias in equity anomaly portfolios

Figure A.I shows the empirical density function (estimated with a gaussian kernel) of the ratio $\sigma(r_{i,t})/\sigma(r_{i,t}^{mean})$ for 430 anomaly portfolios from Haddad et al. (2020) between July 2020 and March 2022. $r_{i,t}$ denotes monthly returns of portfolio i based on common end-of-month index levels. $r_{i,t}^{mean}$ follows the mean-of-month approach as conducted by Chrono24.com, i.e., each month, we calculate the average of daily index levels within a given month and then use them to derive related monthly returns.

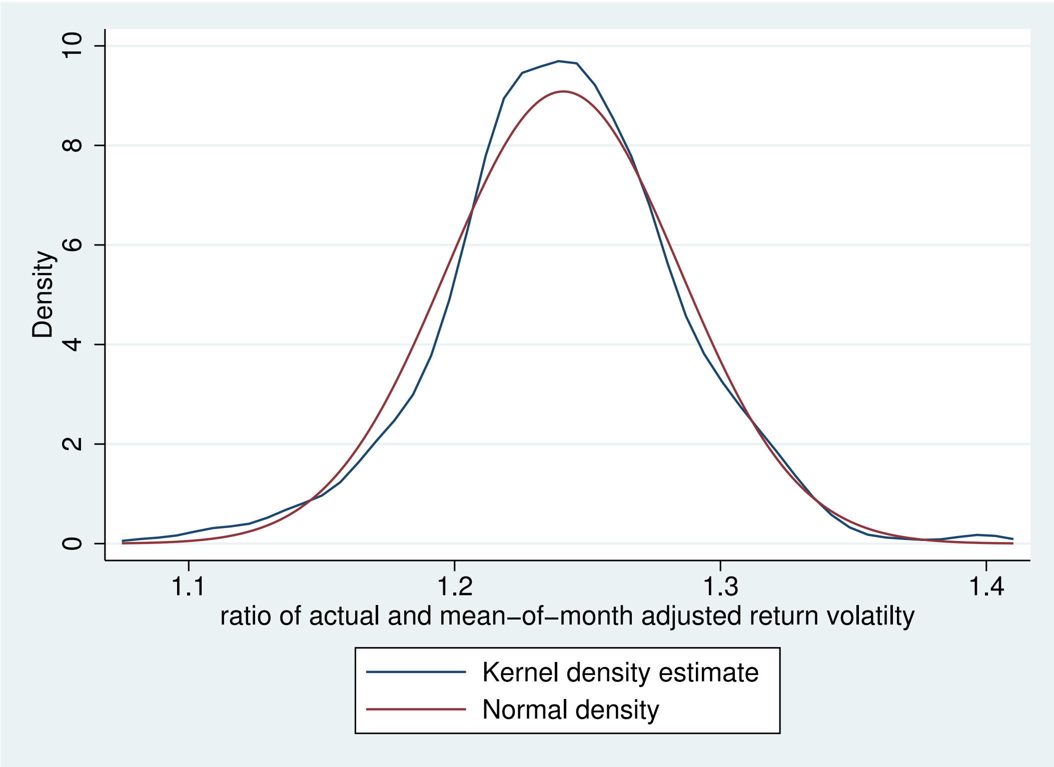

Fig. A.I. The graph shows the empirical density function (estimated with a gaussian kernel) of $\sigma(r_{i,t})/\sigma(r_{i,t}^{mean})$ for 430 anomaly portfolios from Haddad et al. (2020) between July 2020 to March 2022.

We observe an average and median ratio of 1.21, indicating that mean-of-month derived time-series volatility of returns are downwards biased. The minimum ratio of 1.09 is observed for the monthly updated value strategy "valuem" proposed by Asness and Frazzini (2013) and the maximum of 1.40 is found for the industry relative reversal strategy "indrrev" (see Da et al. (2013)).[42]

[42]See Haddad et al. (2020) for strategy definitions.

Volatility scaling vs. advanced ARMA methods

We address the issue of observing luxury watch prices which reflect other than end-of-month values in this section. As described in the main text, our monthly prices represent averages of non-observable daily prices within a given month, which we refer as mean-of-month prices. This induces a downwards bias of estimated time-series volatility for monthly derived luxury watch returns, a problem which is similar with the literature on real estate pricing. There, returns of appraisal-based indices seem to be too smooth, resulting from the appraisal process and temporal aggregation effects in index construction (see Geltner (1993) and Maurer et al. (2004)).

A common approach in the real estate literature to address this problem is to use auto-regressive processes that are fit to the original return series and use related residuals for estimating the "true" return volatility. Compared to our simple volatility-scaling approach described in Section 6.2.2, this requires assumptions about the underlying, unobservable process for daily luxury watch returns which induces further uncertainty to related volatility estimates. For robustness of our approach, we compare our simple volatility-scaling procedure with these more advanced techniques in this section.

For each time-series of monthly luxury watch returns, we follow Firstenberg et al. (1988) and begin with applying an AR(4) model for monthly returns $r_{i,t}$ of luxury watch i:[43]

$$r_{i,t} = \omega_{i,0} + \sum_{p=1}^{4} \omega_{i,p} r_{i,t-p} + \epsilon_{i,t}, \tag{6.4}$$

with $\epsilon \sim \mathcal{N}(0,1)$. Using coefficient estimates, Firstenberg et al. (1988) derives corrected returns $r_{i,t}^{cor.}$:

$$r_{i,t}^{cor.} = \hat{\omega}_{i,0} + \frac{\hat{\epsilon}_{i,t}}{1 - \sum_{p=1}^{4} \hat{\omega}_{i,p}}. \tag{6.5}$$

Subsequently, the mean of AR(4)-model corrected luxury watch returns $r_{i,t}^{cor.}$ equals $\hat{\omega}_{i,0}$ and the time-series volatility equals $\frac{\sigma_{\hat{\epsilon}_{i,t}}}{1-\sum_{p=1}^{4} \hat{\omega}_{i,p}}$.

Figure A.II plots two empirical density function (estimated with a gaussian kernel). First, the blue line shows the distribution of unsmoothed luxury watch returns $r_{i,t}^{cor.}$ following the approach proposed by Firstenberg et al. (1988). Second, the red line shows

[43] A minimum of 12 observations is required for model estimation.

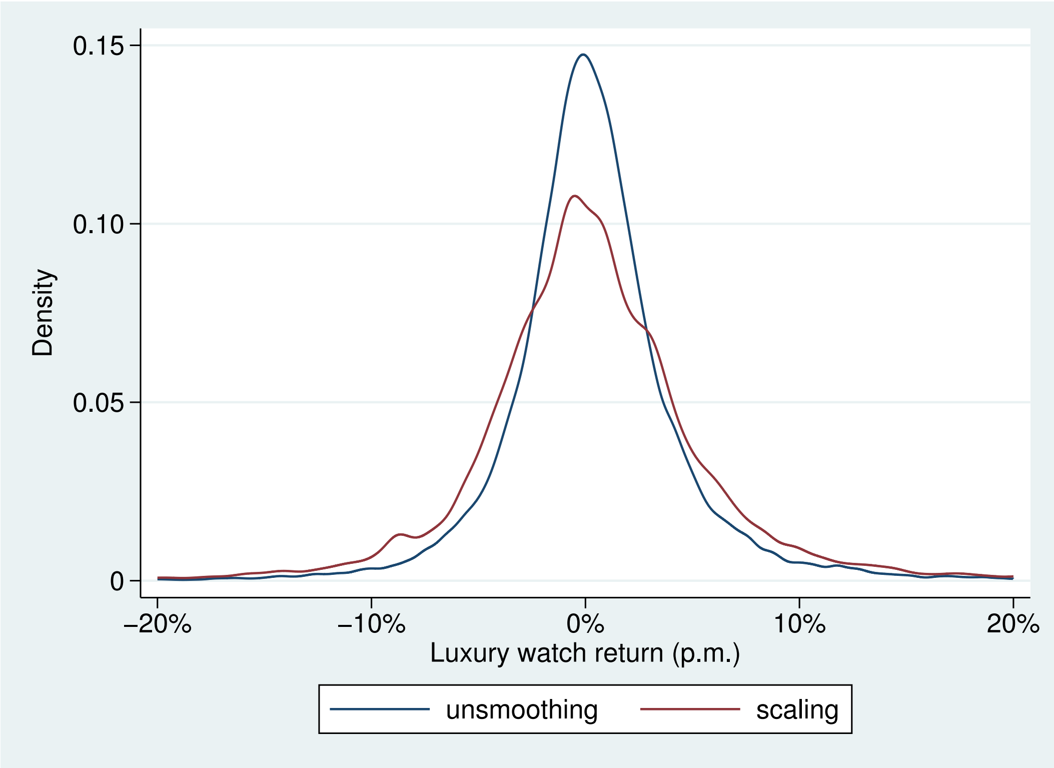

Fig. A.II. This figure shows the empirical density function (estimated with a gaussian kernel) for unsmoothed luxury watch returns (blue line) following the approach proposed by Firstenberg et al. (1988) to account for the downwards bias of volatility in index returns, and for volatility-scaled luxury watch returns (red line) following our simple scaling-approach described in Section 6.2.2.

the distribution of scaled luxury watch returns, following our simple volatility-scaling approach, i.e., we apply the transformation $\bar{r}^w_{i,t} + 1.31 \cdot \left(r^w_{i,t} - \bar{r}^w_{i,t}\right)$ with $\bar{r}^w$ as the vector of demeaned returns for each series i of luxury watch returns $r^w_{i,t}$. The factor of 1.31 corresponds to the 95th percentile of estimated volatility ratios $\frac{\sigma_{r_{i,t}}}{\sigma_{r^{mean}_{i,t}}}$ derived from 430 equity portfolios for 47 anomalies described in Haddad et al. (2020) (see the main text for details).

Unsurprisingly, following the unsmoothing approach proposed in Firstenberg et al. (1988) generates a nearly perfect gaussian distribution of luxury watch returns, induced by auto-regressive model assumptions, whereas our volatility-scaling returns show a higher density in the tails of the distribution. Nevertheless, applying a pooled regression of unsmoothed returns on scaled returns reveals a slope coefficient of 1.0983 with a standard error clustered by watch and month of 0.1570, indicating that our scaling approach sufficiently captures relevant risk properties of the more advanced approaches. Our main results are therefore not affected by choosing the unsmoothing approach from Firstenberg et al. (1988). This is also the case for using the 1-step MA unsmoothing method used for adjusting hedge funds returns as suggested by Getmansky et al. (2004).

A.II Luxury watch returns compared with fine wine and art returns

To put the performance of our price-weighted luxury watch market return $Watch^{Dow}$ in perspective with the performance of other prominent collectibles, we use the Liv-ex Fine Wine 100 Index and the Liquiditiy All Art Index (ARTBnk LLC), both provided by Refinitiv Eikon (Datastream) for our sample period July 2010 to March 2022. The (monthly) wine index tracks the historic prices of 100 wines (mostly Bordeaux and Burgundy) on the world's most active and liquid marketplace, the London International Vintners Exchange. The (annual) index for global fine art is determined by repeat sale pairs of works sold at Christie's, Sotheby's, or Phillips auction houses. We notice that wines had a negative return for our first subperiod until February 2020 of -0.10% per month. Similar with luxury watches, wines strongly performed afterwards with 1.32% on average until March 2022. For the latter subperiod, the according Sharpe ratio is 1.65 indicating a risk-adjusted outperformance of stocks, bonds, commodities, and gold, but clearly underperforming luxury watches. The art market had an average return of 1.75% p.a. between 2010 and 2020 with an according Sharpe ratio of 0.15. However, between 2020 and 2022, we observe an average return of 6.15% p.a. with a Sharpe ratio of 0.92. Together these results provide evidence that fine wine, art and luxury watches outperformed all major asset since February 2020, however, the latter being the most profitable investments.

Chapter 7

Look at my Watch! Continuous Information and the Momentum Effect in the Market for Luxury Watches

This research is joint work with Klaus Röder (University of Regensburg). The paper is currently under review in *Financial Markets and Portfolio Management*. The journal ranking is *B* according to the VHB Publication Media Rating 2024. The paper was presented at the poster session of the 31st Annual Meeting of the German Finance Association (DGF) 2025 in Hagen.

Abstract

We find a consistent momentum return premium of 1.25% p.m. in the global market for luxury watches between June 2017 and September 2024. Investors are inattentive to a series of frequent gradual returns during the momentum formation period and draw more attention to infrequent large returns. For that reason, momentum portfolio profits decrease from 1.67% for luxury watches with continuous information during their formation period to -0.38% for watches with discrete information, but similar cumulative formation-period returns. Conditional on continuous information, momentum profits covary with the market state, return dispersion, and investor sentiment.

Keywords: Empirical Asset Pricing · Alternative Investments · Collectibles · Luxury Watches

JEL classification: G11 · G12 · G15.

7.1 Introduction

The market for luxury watches has experienced rapid growth in recent years. According to the Cap Gemini Wealth Report 2024, ultra-high-net-worth individuals having a net wealth of at least $30 million are reaching unprecedented numbers (220,000 people) and wealth levels. Together with more than 20 million wealthy individuals having at least the status of being millionaire (in U.S.-$), they represent $86,790 billion in total. Their desire to diversify among asset classes drives interest in alternative investments, including commodities, cryptocurrencies, and collectibles. Unsurprisingly, the Deloitte Swiss Watch Industry Study 2023 shows that the Swiss watch industry had an impressive year in 2022 with exports rising to a new all-time record high of CHF 24.8 billion, a 14% increase on 2019's pre-pandemic figures of CHF 21.7 billion.[1]

Despite this great market activity and the increasing interest in luxury watches, they tend to be neglected in the academic literature. The only studies on luxury watches found are Köstlmeier and Röder (2025b) who analyze a broad cross-section of more than 300 luxury watches since 2010 and Köstlmeier and Röder (2025a) who evaluate diversification benefits of luxury watches and day-of-the-week effects. Köstlmeier and Röder (2025b) shows that a price-weighted luxury watch market index gains 0.64% per month and thus outperforms gold, bonds, and commodities on a risk-adjusted basis, but is no hedge or safe haven against them. Their extensive analysis on watch counterparts for prominent stock market return predictors reveals that size, reversal, MAX and momentum form successful long-short strategies. Using a novel data set on luxury watches at the daily level, this study is the first to provide novel insights into the economic channels driving the momentum effect on the market for luxury watches. Our study contributes to the strand of literature that analyzes exotic asset classes like the market for sports betting contracts (Moskowitz (2021)) or the market for cryptocurrencies (Liu et al. (2022)) to shed light onto phenomena on the stock market, with the momentum anomaly as perhaps the most prominent one.

The momentum effect is among the most studied effects in asset pricing and is found not only on the stock market (Jegadeesh and Titman (1993)), but also among commodities (Miffre and Rallis (2007)), corporate bonds (Jostova et al. (2013)), cryptocurrencies (Liu et al. (2022)), currencies and commodity futures (Asness et al. (2013)), and among collectibles like luxury watches (Köstlmeier and Röder (2025b)). The traditional asset pricing literature has proposed a huge variety of theories, both risk and behavioral models,

[1]See Cap Gemini (2024) and Deloitte (2024).

to explain the momentum phenomenon. With respect to the luxury watch momentum premium, our findings are in line with the limited-attention induced underreaction channel proposed in Da et al. (2011), Bae and Wang (2012), and Da et al. (2014).

We consider 124 luxury watch indices of 26 brands from WatchCharts.com and their daily returns for the period 06/30/2017 to 09/30/2024. Similar with other asset classes, we document a strong momentum effect on the market for luxury watches generating a highly significantly value-weighted high-low strategy return of 1.25% per month. We find that the inattentiveness of investors to continuously arriving information during the momentum formation period drives momentum returns. This is known as the frog-in-the-pan hypothesis which originates from limited investor attention. According to the frog-in-the-pan anecdote, a frog will jump out of a pan containing boiling water since the dramatic temperature change induces an immediate reaction. Conversely, if the water in the pan is slowly raised to a boil, the frog will underreact and perish.[2] Using bivariate independent portfolio sorts, we find that momentum returns decrease from a highly significant 1.67% for luxury watches with continuous information during their formation period to an insignificant -0.38% for watches with discrete information, but similar cumulative formation-period returns. Our robustness tests differentiate between information discreteness (Da et al. (2014)), which is motivated by limited attention, and return consistency (Grinblatt and Moskowitz (2004)), whose motivation lies with the disposition effect. Overall, our battery of empirical tests indicates that return consistency and thus the disposition effect is not responsible for the return predictability of continuous information on the market for luxury watches.

We document several additional results. Given that our main findings imply that the luxury watch momentum effect is attributable to underreaction induced mispricing, momentum profits should depend on the state of the market as proposed by Cooper et al. (2004). Conditional on information discreteness, we document that momentum returns are significantly higher by 0.74 percentage points p.m. in up market states compared with down market states. Similar return differences are found conditional on low/high states of return dispersion (Stivers and Sun (2010)) or investor sentiment (Baker and Wurgler (2006)), giving raise to our mispricing hypothesis for the momentum effect. Beyond that, we employ the instrumented principal component analysis (IPCA) introduced in Kelly et al. (2019) to identify latent factors and conditional betas on our comprehensive set of luxury watch characteristics which allows to test implications from our main findings:

[2]No animals were harmed in the writing of this study. The hypothesis is first proposed by Da et al. (2014) who show that the stock momentum effect is much stronger after continuous information.

First, if information discreteness has a conditional relationship with momentum, then we expect both characteristics to enter the same estimated IPCA factor and contribute to this factor to a large extent. Second, if momentum is unlikely to be driven by the disposition effect, we except return consistency to enter an IPCA factor unrelated with momentum. We find that exposure to one of our estimated IPCA factors is mainly driven by return consistency while momentum enters this factor only moderately. Moreover, the factor with highest exposure to momentum is also the factor highly correlated with a low degree of information discreteness. Overall, the luxury watch momentum effect seems to be a phenomenon driven by limited investor attention.

The remainder of this paper is structured as follows. Section 7.2 introduces the data set and displays descriptive statistics for the variables used in this study. Section 7.3 analyzes the momentum effect on the market for luxury watches. Section 7.4 investigates the source of the momentum effect by examining potential economic channels. Section 7.5 presents further robustness tests on the pervasiveness of the luxury watch momentum effect. Section 7.6 concludes and discusses implications for investors, wealth managers, and luxury watch dealers.

7.2 Data

7.2.1 Luxury watches under investigation

We collect daily price data for luxury watches from the website WatchCharts.com. This site is one of the main sources of information on current secondary market prices for new and vintage luxury watches. They collect, structure, and analyze millions of data points to determine daily market prices for more than 27,223 luxury watch models from 337 brands. Original sales data stems from popular platforms and secondary marketplaces around the world, among them e.g., eBay, Reddit, Rolex-Forums, Omega-Forums, ManOnTime, Rakuten (Japan), or Carousell (Asia-Pacific).

In this study, we analyze the returns of luxury watches at the series level.[3] The series indices are well diversified and each of them comprises up to 30 highly liquid tradeable luxury watches from a specific series. First, luxury watches are generally identified by their reference number which is related with specific compositions. This means that

[3] Throughout this study we use the term "model" to refer to luxury watches having the same reference number. The reference number is not unique for each piece, but all watches of a given reference number have exactly the same specifications and are mutually interchangeable. A luxury watch "series" typically comprises multiple watches having different reference numbers to account for different characteristics like the color of bezel or dial and the materials used.

all watches with the same reference number have the exact same specifications as case materials, color of the dial and bezel, or the bracelet. Take e.g., the popular series Rolex Daytona: The discontinued Daytona having a steel case with a white dial has the reference number 116500LN-0001 while the model having a black dial is indicated by the number 116500LN-0002. The currently produced platinum model with blue dial has the reference number 126506-0001. Second, WatchCharts.com aggregates these individual (i.e., per reference number) price data for all watches from a specific series to create a series index. Each index represents an average of 30 models from a given series and is calculated as a weighted average of the (ask) price performance of the set of included watches. In case of the Rolex Daytona series index, the unique Daytona model with white (black) dial comprises a weight of 14.7% (12.4%), and the unique platinum Daytona with reference number 126506-0001 3.6%. To reflect market trends over a long-term period, the set of watches and weights for each index are updated every January 1st.

7.2.2 Descriptive statistics and variables

Our sample comprises a total of 124 luxury watch indices from 26 brands, each of them representing a well diversified, value-weighted index of up to 30 luxury watches of a specific series. Our sample period is 06/30/2017 to 09/30/2024 and all data is denominated in currency U.S.-$. The luxury watch market is dominated by few Swiss manufactures (Statista (2022)). In terms of economic relevance, leading manufactures are Rolex, Patek Philippe, and Audemars Piguet. Looking at Panel A of Table 7.1, our sample reflects this dominance of these three brands among luxury watches as their combined market weight in terms of aggregated prices is 58.22%, although they comprise only 24 out of in total 124 observed luxury watch series.

Table 7.1: Summary statistics of luxury watches.

This table shows summary statistics for the luxury watch series included in our sample. Panel A reports the total number of watch series for each brand and the brand's average percentage weight in terms of aggregated prices. Panel B reports monthly time-series averages of the cross-sectional mean, standard deviation (SD), 5th percentile, 25th percentile, median, 75th percentile, 95th percentile, and the number of observations (N) of global ask prices for luxury watches. Panel C shows averages of monthly value-weighted (VW) and equal-weighted (EW) average luxury watch returns, and monthly cross-section standard deviations (SD) of returns for all luxury watches (Market), for micro, small, and big watches, and for the brands Patek Philippe, Rolex, and Audemars Piguet. The last two rows refer to watches having a price below (above) the 20th (80th) percentile of prices. Luxury watches in the top 90% of aggregated market prices are classified as big. Small watches are between the 3rd and 10th percentile and watches below the 3rd percentile are micro. We assign watches to these size groups at the end of each month. The sample period is June 2017 to September 2024 and all data is denominated in U.S.-$.

Panel A: Luxury watch brands ($N = 124$)

Brand	Series	Weight	Brand	Series	Weight
Patek Philippe	7	25.85	Omega	8	4.21
Rolex	13	20.70	IWC	6	3.66
Audemars Piguet	4	11.67	Cartier	10	3.22
Vacheron Constantin	5	5.07	Breitling	15	2.97
A. Lange & Söhne	3	4.48	Other	53	18.17

Panel B: Luxury watch prices (in U.S.-$)

	Mean	SD	5th	25th	Median	75th	95th	N
Market	14,980	23,988	1,396	4,179	7,602	13,636	49,680	76
Micro	2,057	960	1,130	1,343	1,758	2,773	3,416	17
Small	4,722	511	3,980	4,366	4,735	5,093	5,414	17
Big	23,661	29,089	6,441	8,642	11,819	26,668	71,967	43
Patek Philippe	81,396	65,110	34,682	40,346	68,463	108,852	157,561	4
Rolex	21,022	17,040	4,441	7,783	15,553	34,842	46,430	10
Audemars Piguet	51,800	25,970	34,639	39,060	44,274	60,168	75,080	2
Other	7,044	5,695	1,279	3,610	5,238	9,176	17,018	59
Low ($< P_{20\%}$)	2,087	954	1,135	1,360	1,832	2,773	3,446	16
High ($> P_{80\%}$)	47,905	36,273	21,776	26,417	36,710	50,821	118,308	16

Panel C: Luxury watch returns (in %)

	N	Percent of Agg. Market Cap.	VW Average return		EW Average return		Cross-section SD of returns
			Mean	SD	Mean	SD	
Market	76	100.00	0.51	1.41	0.26	0.69	1.32
Micro	17	2.77	0.19	0.72	0.16	0.67	1.32
Small	17	6.77	0.09	0.64	0.10	0.65	1.22
Big	43	90.45	0.56	1.52	0.35	0.85	1.32
Patek Philippe	4	27.92	0.64	2.70	0.50	1.80	1.78
Rolex	10	24.29	0.61	1.59	0.58	1.44	1.38
Audemars Piguet	2	13.70	0.84	2.36	0.74	1.91	1.30
Other	59	34.10	0.03	0.53	0.04	0.52	1.13
Low ($< P_{20\%}$)	16	2.91	0.21	0.75	0.15	0.67	1.33
High ($> P_{80\%}$)	16	64.89	0.71	1.91	0.59	1.16	1.17

Table 7.1 presents summary statistics for the cross-section of monthly luxury watch series prices (Panel B) and returns (Panel C). In line with the approach in Fama and French (2008b) for the stock market, we classify luxury watches in the top 90% of aggregated market price as big and watches in the bottom 3% are classified as micro. Luxury watches having a price between the 3rd and 10th percentile of aggregated market price are denoted as small. We assign watches to these three size groups at the end of each month. For perspective, the average price breakpoint separating micro from small and small from big are \$3,942 and \$5,665, respectively. There is, however, no clear definition for the term luxury watch and most of definitions refer to qualitative attributes such as exclusivity or brand status rather than price categories (Lazazzera (2023)). The Federation of the Swiss Watch Industry FH reports that Swiss watch exports amount to 26.7 billion CHF in 2023, including all types of watches and exceeding 2022 export value by 7.6%. What stands out is that watches with a price below 500 CHF represented more than 80% of the increase in the total number of exported items. At the other end of the price scale, 92% of the growth in export value is generated by watches priced at over 3,000 CHF. The average minimum watch price in our sample exceeds \$1,000 and all series are watches with (non-electronic) mechanical movements, so we conclude that all watches in our sample fulfill the definition of luxury watches. Further, our breakpoints separating micro from small luxury watches seem to be reasonable given the 3,000 CHF (ca. \$3,330) cutoff reported in Federation of the Swiss Watch Industry FH (2024).

As shown in Panel B of Table 7.1, the average luxury watch in our sample costs \$14.980 while the median price is \$7,602. 90% of all watches under study are within the price range between \$1,396 and \$49,680. Looking at the brand level, we notice that watches produced by Patek Philippe have a much higher average price of \$81,396 compared to Audemars Piguet (\$51,800) or Rolex (\$21,022). Panel C reveals that we observe 76 luxury watch series on average and almost half of them are micro or small, despite only accounting for less than 10% of aggregated market prices. While the distribution across size segments is not as extreme compared with stocks (see Fama and French (2008b)), the numerous micro and small watches influence the watch market return. The equal-weighted market return is only 0.26% p.m. whereas the value-weighted market return generates 0.51%. This is the result of the high (equal-weighted) return of 0.35% of big watches. The portfolios comprising watches by Patek Philippe, Rolex, or Audemars Piguet highlight the fact that most expensive watches outperform less expensive ones. The average value-weighted portfolio return for these brands exceeds 0.61% p.m. while all other brands only yield 0.03%. In contrast with stocks, big watches are the ones having a higher volatility compared with small and micro watches, both in the time-series and the cross-section of returns. Nevertheless, volatilities are remarkable low compared to stocks. The average standard deviation of the monthly return on the value-weighted U.S. stock market (4.85%) during our sample period far exceeds the time-series volatility for the luxury watch market return of only 1.41%.[4] Although the U.S. stock market generated an average value-weighted return of 1.25% p.m. from Jan. 2017 to Sept. 2024 compared with only 0.51% for luxury watches, annualized Sharpe ratios are essentially identical (luxury watches: 0.78, U.S. stocks: 0.77).[5]

Our set of chosen variables follows Köstlmeier and Röder (2025b) who consider a comprehensive list of well-established equity strategies from Feng et al. (2020) and Chen and Zimmermann (2022) and select characteristics that can be constructed using price information of luxury watches. The analysis in Köstlmeier and Röder (2025b) is, however, limited to a monthly periodicity. In lack of daily data, they choose variable definitions that best suit common definitions for the equity market. This study fills this gap, and the construction of variables closely follows common procedure for stocks. The variables used in this study are defined as follows: A watch's size (*SIZE*) is defined as the natural log of the market price at the end of the previous month. *BETA* is the slope coefficient from a time-series regression of a watch's excess return on the excess return of the (watch) market portfolio using one month of daily return data. *SKEW* is total skewness using daily watch returns over the previous month and *IVOL* is the annualized idiosyncratic volatility relative to the (watch) market model using daily watch returns over the previous month.[6] Following Jegadeesh and Titman (1993) and Köstlmeier and Röder (2025b), momentum (*MOM*) is the cumulative prior four-month watch return, skipping the most recent month. Short-term reversal (*REV*) is the monthly watch return over the previous month (Jegadeesh (1990)). Following Bali et al. (2011), *MAX* is the maximum daily watch return over the previous month. As in Da et al. (2014), information discreteness *ID* is defined as $\text{sgn}(MOM) \times [\%neg - \%pos]$ where $\%pos$ and $\%neg$ denote the respective percentage of positive and negative daily returns during the formation period of *MOM*. Return consistency (*RC*) is a dummy variable equaling one if a watch's monthly returns are positive (negative) for all months during the formation period of *MOM* and zero otherwise (Grinblatt and Moskowitz (2004)). In line with Da et al. (2011), *SVI* is the

[4] Data for the U.S. stock market is retrieved from K. French data library.

[5] We use one-month Treasury Bill rates from K. French data library as a proxy for the risk-free rate.

[6] A minimum of 21 observations per month is required for the calculation of *BETA*, *SKEW*, and *IVOL*. In contrast to equities which can only be traded on five business days, luxury watches can be traded throughout each day of the week.

natural log of the Google search volume index at the end of the previous month where the combination of brand name and luxury watch series name is used for the search queries. $ASVI$ is the ratio of SVI and the average SVI over the past three months.

Table 7.2: Summary statistics of variables.

This table shows summary statistics for the variables used in this study. $SIZE$ is the natural log of market price at the end of the previous month. $BETA$ is the slope coefficient from a time-series regression of the watch's excess return on the excess return of the (watch) market portfolio using one month of daily return data. MOM is the cumulative prior four-month watch return, skipping the most recent month. Short-term reversal (REV) is the monthly watch return over the previous month. $SKEW$ is total skewness using daily watch returns over the previous month and $IVOL$ is the annualized idiosyncratic volatility relative to the (watch) market model using daily watch returns over the previous month. MAX is the maximum daily watch return over the previous month. Information discreteness ID is defined as sgn(MOM) $\times$ [$\%neg - \%pos$] where $\%pos$ and $\%neg$ denote the respective percentage of positive and negative daily returns during the formation period of MOM. Return consistency (RC) is a dummy variable equaling one if a watch's monthly returns are positive (negative) for all months during the formation period of MOM and zero otherwise. SVI is the natural log of the Google search volume index at the end of the previous month where the combination of brand name and luxury watch series name is used for the search queries. $ASVI$ is the ratio of SVI and the average SVI over the past three months. The sample period is June 2017 to September 2024.

Variable	Mean	SD	5th	25th	Median	75th	95th
$SIZE$	8.98	1.07	7.35	8.36	8.93	9.47	10.79
$BETA$	0.43	1.43	−0.93	−0.18	0.15	0.68	2.38
MOM	0.76	2.98	−2.91	−0.97	0.35	2.02	5.63
REV	0.24	1.39	−1.62	−0.48	0.10	0.85	2.44
$SKEW$	0.05	1.90	−2.95	−1.03	0.15	1.12	2.97
$IVOL$	2.67	2.08	0.67	1.31	2.12	3.38	6.16
MAX	0.48	0.47	0.08	0.19	0.35	0.61	1.28
ID	−0.11	0.13	−0.35	−0.17	−0.09	−0.03	0.03
RC	0.44	0.49	0.00	0.00	0.29	0.95	1.00
SVI	3.99	0.39	3.30	3.84	4.08	4.26	4.42
$ASVI$	0.02	0.20	−0.20	−0.06	0.00	0.07	0.31

Table 7.2 summarizes the main variables in our study. Focusing on the key variable of interest, MOM, we observe that a typical watch in our sample has a cumulative four-month return of 0.76% with a cross-sectional standard deviation of 2.98%. The average ID is a negative -0.11% indicating that watch returns tend to be quite consistent, in line with our previous results that luxury watch returns have a very low volatility. Looking at RC, 44% of luxury watches have only positive or negative returns in all month of the momentum formation period. Table 7.3 shows cross-sectional correlations between all variables and one-month ahead luxury watch series returns.

Table 7.3: Correlation coefficients of variables and future watch returns.

This table contains the cross-sectional correlations between the variables used in this study and one-month ahead luxury watch returns RET_{t+1}. Below-diagonal entries represent Pearson correlations and above-diagonal entries the average Spearman rank correlation. The sample period is June 2017 to September 2024.

	RET_{t+1}	$SIZE$	$BETA$	MOM	REV	$SKEW$	$IVOL$	MAX	ID	RC	SVI	$ASVI$
RET_{t+1}	1.00	0.05	0.09	0.31	0.27	0.12	0.11	0.18	−0.10	0.05	−0.02	0.05
$SIZE$	0.08	1.00	0.22	0.09	0.07	0.05	−0.24	−0.16	−0.28	0.22	−0.06	0.04
$BETA$	0.13	0.31	1.00	0.13	0.09	0.06	0.24	0.23	−0.11	0.05	0.02	0.01
MOM	0.28	0.14	0.18	1.00	0.32	0.14	0.16	0.22	−0.15	0.09	0.01	0.01
REV	0.25	0.09	0.14	0.30	1.00	0.56	0.12	0.42	−0.12	0.06	−0.01	0.04
$SKEW$	0.11	0.04	0.05	0.12	0.46	1.00	0.06	0.46	−0.06	0.03	−0.03	0.03
$IVOL$	0.10	−0.18	0.28	0.13	0.12	0.05	1.00	0.83	0.02	−0.09	0.09	−0.03
MAX	0.14	−0.07	0.37	0.18	0.41	0.38	0.84	1.00	−0.02	−0.06	0.05	0.00
ID	−0.13	−0.34	−0.13	−0.24	−0.16	−0.08	0.05	0.01	1.00	−0.52	−0.13	0.01
RC	0.05	0.23	0.05	0.12	0.07	0.03	−0.10	−0.06	−0.50	1.00	0.06	0.00
SVI	0.02	−0.01	0.05	0.04	0.03	−0.01	0.11	0.08	−0.14	0.06	1.00	0.16
$ASVI$	0.01	0.02	−0.01	−0.01	0.01	0.01	−0.04	−0.03	0.02	−0.02	0.14	1.00

What stands out is the moderate positive correlation between one-month ahead returns and momentum (0.28) and reversal (0.25). The high correlation between *MAX*, *IVOL*, and *REV* is a mechanical relationship. Interestingly, *ID* and *SIZE* are negatively correlated which implies that most expensive watches have more consistent returns.

7.3 Momentum and subsequent luxury watch series returns

To test whether there is a momentum effect in luxury watch series markets, we examine the relation between momentum and future returns using the methodology presented in Fama and MacBeth (1973). In particular, we estimate cross-sectional regressions of monthly luxury watch returns $r_{i,t+1}$ for series i on momentum (*MOM*) and common control variables. We estimate different variations of the following cross-sectional regression with full specification:

$$r_{i,t+1} = \alpha_{0,t} + \beta_{0,t} MOM_{i,t} + \sum_{j=1}^{N} \beta_{j,t} X_{i,t} + \epsilon_{i,t+1}. \tag{7.1}$$

Because *MOM* is related to the variables presented in the previous section, we have to control for them to ensure that the momentum effect is robust. For that reason, the vector X holds these control variables and our first specifications use each of them as the sole return predictor to measure its unique strength. OLS-regressions imply equal weights on all observations, thus emphasizing inexpensive luxury watches. To address this issue, all regressions are estimated using a value-weighted (VW) scheme by default, i.e., they use the natural log of the watch price (*SIZE*) as observation weight. For robustness, our full regression specification is also estimated using the traditional equal-weighting (EQ) scheme.

Table 7.4: Cross-sectional regressions of monthly watch returns.

This table shows time-series averages of monthly estimates from cross-sectional regressions including an intercept α_0. The dependent variable is the luxury watch series return RET_{t+1} and the explanatory variables as of month t are given in the first column. $SIZE$ is the natural log of market price at the end of the previous month. $BETA$ is the slope coefficient from a time-series regression of the watch's excess return on the excess return of the (watch) market portfolio using one month of daily return data. MOM is the cumulative prior four-month watch return, skipping the most recent month. Short-term reversal (REV) is the monthly watch return over the previous month. $SKEW$ is total skewness using daily watch returns over the previous month and $IVOL$ is the annualized idiosyncratic volatility relative to the (watch) market model using daily watch returns over the previous month. MAX is the maximum daily watch return over the previous month. ID is defined as $\text{sgn}(MOM) \times [\%neg - \%pos]$ where $\%pos$ and $\%neg$ denote the respective percentage of positive and negative daily returns during the formation period of MOM. Return consistency (RC) is a dummy variable equaling one if a watch's monthly returns are positive (negative) for all months during the formation period of MOM. SVI is the natural log of the Google search volume index at the end of the previous month where the combination of brand name and luxury watch series name is used for the search queries. $ASVI$ is the ratio of SVI and the average SVI over past three months. The R^2 values are adjusted for degrees of freedom. N denotes the average number of luxury watch series. $Weighting$ indicates if a value-weighted (VW) or equal-weighted (EQ) weighting scheme is used in the cross-sectional regression approach. Newey and West (1987) corrected t-statistics using a lag of six are given in parenthesis and */**/*** indicate significance at the 10%/5%/1% level. The sample period is June 2017 to September 2024.

Specification	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
α_0	−1.57**	0.00	−0.06	−0.04	0.09	−0.08	−0.12	−0.10	0.00	−0.45	0.09	0.15	0.10
	(−2.33)	(0.00)	(−0.47)	(−0.26)	(0.48)	(−0.52)	(−0.81)	(−0.76)	(0.01)	(−1.52)	(0.46)	(0.22)	(0.14)
$SIZE$	0.18**											−0.01	−0.00
	(1.99)											(−0.21)	(−0.04)
$BETA$		0.19***										0.13***	0.12***
		(2.87)										(2.86)	(2.79)
MOM			0.14***									0.07***	0.07***
			(6.30)									(3.49)	(3.77)
REV				0.26***								0.18*	0.17*
				(4.34)								(1.74)	(1.77)
$SKEW$					0.07***							0.01	0.01
					(3.88)							(0.65)	(0.86)
$IVOL$						0.09**						0.07	0.08
						(2.32)						(1.16)	(1.17)
MAX							0.51***					−0.57	−0.58
							(3.87)					(−1.16)	(−1.21)
ID								−1.36**				−0.91**	−0.87**
								(−2.01)				(−2.25)	(−2.12)
RC									0.19*			−0.08	−0.09
									(1.78)			(−1.16)	(−1.23)
SVI										0.14		−0.04	−0.05
										(1.33)		(−0.68)	(−0.87)
$ASVI$											0.14	0.19	0.19
											(0.88)	(0.93)	(0.98)
$Adj.R^2$	0.06	0.08	0.13	0.12	0.02	0.05	0.06	0.09	0.01	0.00	0.00	0.26	0.28
N	69	69	69	69	69	69	69	69	69	69	69	69	69
$Weighting$	VW	VW	VW	VW	VW	VW	VW	VW	VW	VW	VW	EQ	VW

Specifications (1) to (11) of Table 7.4 presents average coefficient estimates from monthly cross-sectional regressions of monthly luxury watch returns on each variable introduced in the previous section. We find highly significant coefficient estimates for *BETA*, *MOM*, *REV*, *SKEW*, and *MAX*. As indicated by the sign of their coefficient estimates, these variables are positively related with one-month ahead luxury watch series returns. The positive relation with *BETA* implies a positive risk-return relationship, thus investors show a risk-averse behavior. The positive relation between *REV* and one-month ahead returns is in line with findings in Köstlmeier and Röder (2025b) who document a strong negative reversal effect until Feb. 2020 and a strong inverted reversal effect (i.e., a positive relation) since March 2020. To our surprise, we find that high *MAX* luxury watches seem to generate significantly higher future returns than low *MAX* watches, contradicting findings for the equity market (see Bali et al. (2011) and Walkshäusl (2014)). However, watches with extreme positive daily returns naturally induce return volatility in the same month, so high *MAX* watches are very likely also watches with high volatility. As shown in Table 7.3, the average correlation between *MAX* and *IVOL* is 0.84. Once we control for this effect in our full specifications, the puzzling relation disappears.

Focusing on momentum, we find a highly significant positive coefficient estimate of 0.14 when *MOM* is the sole return predictor in the cross-section of luxury watch series returns. When *MOM* is combined with common control variables as shown in specifications (12) and (13), we observe that momentum is not explained away in presence of all other predictor variables. When the full set of control variables is applied, the average *MOM* coefficient estimate amounts to 0.07 in both equal- and value-weighted regressions, with a robust t-statistic exceeding 3.49. Although the economic effect of *MOM* in the full specification is only half of the univariate specification, the presence of information discreteness *ID* renders the lottery-like characteristics *SKEW*, *IVOL*, and *MAX* insignificant.

To assess the economic return effect of MOM in the spirit of a long-short decile portfolio sort investing in high *MOM* luxury watches and short-selling low *MOM* luxury watches we use the average difference between the 10th and 90th percentiles of *MOM*, which are -2.19 and 4.10, respectively. In this way, the average *MOM* coefficient of 0.07 is equivalent to an average return premium of 0.44% p.m. (formally, $0.07 \cdot (4.10 - (-2.19))$). Similarly, if we multiply the estimated average slope coefficient on *MOM* with the average cross-sectional standard deviation of *MOM* (2.98, see Table 7.2), we find that a one-standard deviation increase in *MOM* is associated with a 0.21% p.m. difference in expected returns. High *MOM* is also associated with low *ID* as indicated by the negative correlation coefficient of -0.24, so both effects may reinforce each other. Using the 10th (-0.27) and 90th (0.01) percentiles of *ID* in a similar manner for a long-short decile portfolio strategy, we observe a return premium of 0.24% p.m. (formally, $-0.87 \cdot (-0.27 - 0.01)$) when taking a long position in the low *ID* decile of luxury watches and selling watches belonging to the highest decile of *ID*.

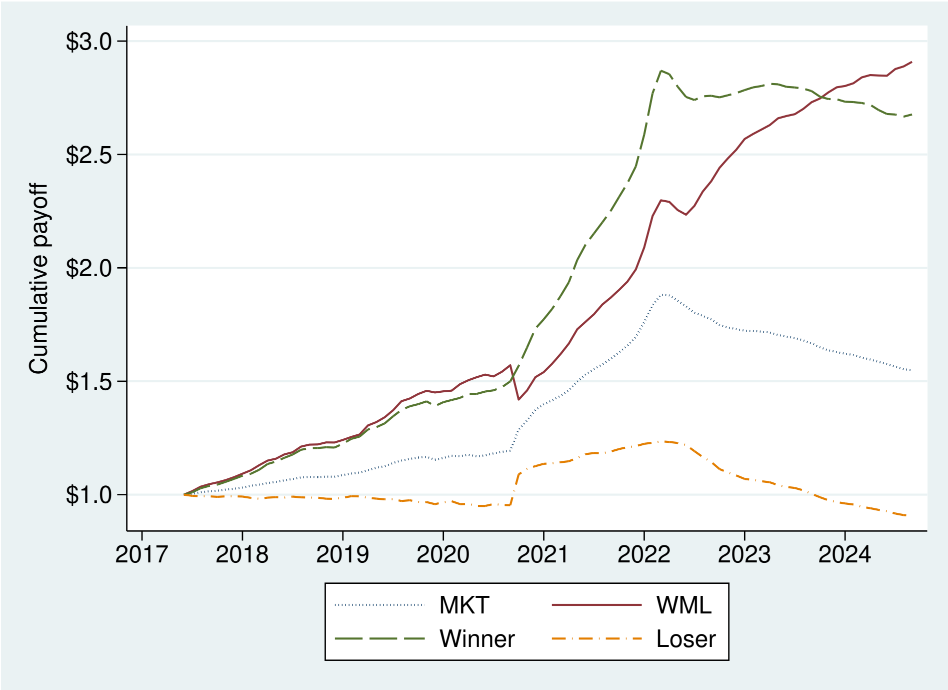

Fig. 7.1. This figure shows the cumulative monthly long-short portfolio returns for the momentum strategy WML (solid line) and the value-weighted luxury watch market portfolio MKT (dotted line). For each month t, the long-short returns are calculated as the difference between top 30% and bottom 30% value-weighted portfolio returns. Luxury watches are allocated to these portfolios based on MOM at the end of each month $t-1$. MOM is the cumulative watch return over formation months $t-4$ to $t-1$. The sample period is June 2017 to September 2024.

To visualize the profitability of the momentum strategy in the market for luxury watches, Figure 7.1 shows the cumulative payoff of an initial $1.00 investment into a strategy that is long in high MOM watches and short in low MOM watches. To be specific, we allocate watches into three portfolios based on MOM using the 30th and 70th percentiles as breakpoints. The long-short returns WML are calculated as the difference between high and low value-weighted portfolio returns. For comparison, the dotted line shows the cumulative payoff of the value-weighted luxury watch market portfolio return. The market portfolio steadily increased to $1.19 until September 2020 and surged to $1.88 in March 2022, whereas the momentum portfolio already reached $2.30. Since then, the market portfolio decreased to $1.55 until September 2024 whereas the momentum portfolio continued its increase to a terminal value of $2.91. To mitigate concerns that short-selling luxury watches is not feasible, Figure 7.1 also plots the cumulative payoff of the long momentum leg (winner) and short leg (loser) separately. Clearly, long-short momentum gains are generated by the respective long leg, i.e. high momentum watches, which generate a cumulative payoff of \$2.68. Investing in past loser watches results in a terminal value of \$0.91 for each dollar invested and ignoring them does not harm our overall momentum portfolio strategy.

7.4 Momentum and information discreteness

In this section, we investigate how *MOM* is related to continuous information in the market for luxury watches. Previous works on the momentum effect in equity markets find a strong relation between momentum and return consistency. For instance, Grinblatt and Han (2005), Watkins (2003), and Grinblatt and Moskowitz (2004) find strong evidence that winner consistency is important. Achieving a high past return with a series of steady positive months appears to generate a larger expected return than a high past return achieved with just a few extraordinary months. Despite the similar construction of *ID* and *RC*, the economic motivation underlying information discreteness differs considerably from return consistency, since *ID* is based on limited attention, while return consistency is based on the disposition effect.

To distinguish between the economic implications of these two channels in a preliminary analysis, we regress *MOM* on lagged values of *ID* or *RC* in univariate regressions to examine their ability to predict momentum. From these (untabulated) regressions, we obtain a significantly negative coefficient estimate on *ID* (t-statistic: -2.39) and a significantly positive estimate on *RC* (t-statistic: 2.12). To explore the return effect associated with momentum among different levels of information discreteness and return consistency, we proceed as follows.

We estimate cross-sectional regressions of monthly luxury watch series returns on lagged *MOM* along with the full set of control variables as outlined in regression Eq. (7.1) in Section 7.3. We add to this regression an interaction term between *MOM* and *ID* or *RC*.Based on our previous findings, we expect a negative estimate on the interaction term between *MOM* and information discreteness, and a positive estimate on the interaction term between *MOM* and return consistency.

Table 7.5: Cross-sectional regressions of monthly watch returns on momentum and common controls.

This table shows time-series averages of monthly estimates from cross-sectional regressions of monthly luxury watch returns on an intercept α_0, momentum (MOM), and common controls. The set of control variables includes watch size ($SIZE$), market beta ($BETA$), reversal (REV), skewness ($SKEW$), idiosyncratic volatility ($IVOL$), the maximum daily return over the previous month (MAX), Google search volume index (SVI), abnormal Google search volume index ($ASVI$), information discreteness (ID), and return consistency (RC). DID is a dummy variable that takes the value one if ID is below the median ID in that month and zero otherwise. Similarly, $DMOM$ and DRC are dummy variables taking the value one if MOM, resp. RC, is above the median MOM (RC) in that month and zero otherwise. $posID$ and $negID$ are signed versions of ID. $posID$ equals $\%pos - \%neg$ if $MOM > 0$ and zero otherwise. $negID$ equals $\%neg - \%pos$ if $MOM < 0$ and zero otherwise. $\%pos$ and $\%neg$ denote the respective percentage of positive and negative daily returns during the formation period of MOM. Similarly, $posRC$ and $negRC$ refer to positive and negative RC dummy variables, respectively. $ID_{\perp RC}$ ($RC_{\perp ID}$) is the portion of ID (RC) that is orthogonal to RC (ID), i.e., it is the sum of intercept and residual from monthly cross-sectional regressions. The R^2 values are adjusted for degrees of freedom. N denotes the average number of luxury watch series. $Weighting$ indicates if a value-weighted (VW) or equal-weighted (EQ) weighting scheme is used in the cross-sectional regression approach. Newey and West (1987) corrected t-statistics using a lag of six are given in parenthesis and */**/*** indicate significance at the 10%/5%/1% level. The sample period is June 2017 to September 2024.

Specification	(1)	(2)	(3)	(4)	(5)	(6)	(7)
α_0	0.10	−0.01	0.15	0.56	−0.38	0.10	0.10
	(0.14)	(−0.02)	(0.19)	(0.65)	(−0.63)	(0.14)	(0.14)
$SIZE$	−0.02	0.01	−0.01	−0.05	0.01	−0.00	−0.00
	(−0.22)	(0.09)	(−0.14)	(−0.52)	(0.12)	(−0.04)	(−0.04)
$BETA$	0.12***	0.10***	0.15***	0.07	0.13**	0.12***	0.12***
	(2.98)	(2.38)	(2.70)	(1.39)	(2.54)	(2.79)	(2.79)
REV	0.18*	0.02	−0.05	0.13	0.26	0.17*	0.17*
	(1.94)	(0.26)	(−0.38)	(1.10)	(1.58)	(1.77)	(1.77)
$SKEW$	0.02	0.01	0.01	0.01	0.01	0.01	0.01
	(1.15)	(0.40)	(0.33)	(0.26)	(0.36)	(0.86)	(0.86)
$IVOL$	0.07	0.04	−0.02	0.06	0.07	0.08	0.08
	(1.09)	(0.53)	(−0.15)	(0.95)	(0.93)	(1.17)	(1.17)
MAX	−0.58	−0.06	0.17	−0.35	−0.71	−0.58	−0.58
	(−1.33)	(−0.19)	(0.28)	(−0.96)	(−1.11)	(−1.21)	(−1.21)
SVI	−0.03	−0.05	−0.05	−0.07	0.05	−0.05	−0.05
	(−0.43)	(−0.72)	(−0.60)	(−1.03)	(0.68)	(−0.87)	(−0.87)
$ASVI$	0.16	0.51	0.58	0.32	−0.07	0.19	0.19
	(0.74)	(1.63)	(1.56)	(1.24)	(−0.27)	(0.98)	(0.98)
MOM	−0.00	0.05*	0.01	0.02	−0.03	0.07***	0.07***
	(−0.04)	(1.92)	(0.33)	(1.22)	(−0.50)	(3.77)	(3.77)
$MOM \times ID$	−0.46***		−0.54***				
	(−3.29)		(−3.23)				
$MOM \times RC$		0.07***	0.02				
		(2.79)	(0.74)				
$DMOM \times DID$				0.40**			
				(2.19)			
$DMOM \times DRC$				0.21			
				(1.52)			
ID	−0.58	−0.64*	−0.13	0.05			−0.67*
	(−1.61)	(−1.73)	(−0.31)	(0.12)			(−1.88)
$ID_{\perp RC}$						−0.87**	
						(−2.12)	
$posID$					2.68***		
					(3.87)		
$negID$					−1.70**		
					(−2.49)		
RC	−0.03	−0.09	−0.02	−0.03		0.06	
	(−0.54)	(−1.06)	(−0.24)	(−0.30)		(0.89)	
$RC_{\perp ID}$							−0.09
							(−1.23)
$posRC$					−0.04		
					(−0.36)		
$negRC$					−0.03		
					(−0.33)		
$Adj.R^2$	0.29	0.30	0.32	0.26	0.30	0.28	0.28
N	69	69	69	69	69	69	69
$Weighting$	VW	VW	VW	VW	VW	VW	VW

Table 7.5 shows average coefficient estimates from the outlined monthly cross-sectional, value-weighted regressions. First, specifications (1) and (2) provide evidence that both *ID* and *RC* add to the explanation of the cross-section of luxury watch series returns. The signs of the interaction terms are in line with our former findings. These findings suggest that the momentum return effect is significantly more pronounced among watches with low information discreteness or high return consistency. Adding both interaction terms in our regression, specification (3) implies that information discreteness is the true effect, for which return consistency is just a proxy. Another key finding is that the estimated coefficient for *MOM* is insignificant in all specifications (1) to (3). This implies that *MOM* alone is not generating significant differences in expected luxury watch returns, but only the consideration of *ID* results in expected return differences.

Because interaction coefficient estimates are hard to interpret and to assess the economic relevance of the effect, we re-estimate regression specification (3) from Table 7.5 using interaction terms between momentum dummy variables and corresponding *ID*, resp. *RC*, dummy variables. If a luxury watch's *ID* (*RC*) is below (above) the median *ID* (*RC*), then the corresponding dummy variable takes the value one and zero otherwise. The dummy variable for *MOM* is defined in the same way. Thus, the average coefficient estimate on the dummy interaction terms provides the economic difference of the momentum effect between luxury watches with high and low levels of *ID*, resp. *RC*, using the median for separating watches into high and low groups. Specification (4) shows the resulting coefficient estimates. First, the coefficient estimate for *MOM* is close to zero, so we find no overall momentum effect once we control for different levels of momentum and information discreteness. The coefficient for the interaction term $DMOM \times DID$ implies that only above median *MOM* and below medium *ID* watches carry a return premium of 0.40% p.m.

To further distinguish between the economic implications of *ID* and *RC*, we examine their respective impacts on past winners and past losers separately. Using signed versions for both variables in specification (5), we conclude that limited attention in the meaning of Da et al. (2014) explains the return continuation of both past winners and past losers. Both, the positive coefficient estimates for *posID* and the negative coefficient estimate for *negID* are statistically significant and economically large in magnitude.

Finally, we separate the effect of *ID* and *RC* in specification (6) and (7) using orthogonalization as a robustness test. Specification (7) demonstrates that the portion of *ID* which is orthogonal on *RC* carries an important information about expected returns as indicated by the significant coefficient estimate of -0.87. On the other hand, specification (7) shows that return consistency left unexplained by *ID* is statistically insignificant and the coefficient estimate close to zero.

Table 7.6: Bivariate independent portfolio analysis of MOM and ID.

This table presents average subsequent value-weighted returns from bivariate independent portfolio sorts. At the end of each month, each luxury watch series is independently allocated to three groups of momentum (MOM) and three groups of information discreteness (ID) based on 30th and 70th percentiles as breakpoints. MOM is the cumulative prior four-month watch return, skipping the most recent month. ID is defined as sgn(MOM) $\times$ [$\%neg - \%pos$] where $\%pos$ and $\%neg$ denote the respective percentage of positive and negative daily returns during the formation period of MOM. Alphas according to the value-weighted watch market return (CAPM) are shown in brackets. Newey and West (1987) corrected t-statistics using a lag of six are given in parenthesis and */**/*** indicate significance at the 10%/5%/1% level. The sample period is June 2017 to September 2024.

	Information Discreteness ID		
	low	mid	high
low MOM	-0.42^{**}	-0.04	0.47
	$[-0.58]^{***}$	$[-0.23]^{***}$	$[-0.24]$
mid MOM	0.20	0.16	0.04
	$[-0.01]$	$[-0.00]$	$[-0.07]$
high MOM	1.26^{***}	0.69^{***}	0.09
	$[0.78]^{***}$	$[0.35]^{***}$	$[-0.04]$
high - low	1.67^{***}	0.74^{***}	-0.38
	$[1.35]^{***}$	$[0.58]^{***}$	$[0.21]$
t(high - low)	7.29	4.61	-1.12

To provide further evidence for robustness, Table 7.6 presents value-weighted average monthly returns for portfolios sorted independently on *MOM* and *ID*. At the end of each month, watches are independently allocated to three portfolios based on the 30th and 70th percentiles breakpoints for each variable. We calculate value-weighted returns for the resulting nine portfolios for the subsequent month. Our results are in line with Da et al. (2014), so momentum profits are significantly larger when *ID* is low compared to when it is high. While the momentum return spread is a large 1.67% p.m. for watches with continuous information (i.e., low *ID*), the momentum effect is not prevalent among watches with rather discrete information as shown by the insignificant return spread of -0.38%. Consequently, continuous information induces stronger and more persistent return continuation than discrete information and luxury watch investors seem to underreact towards continuous information. These results are obviously robust to short-selling restrictions: Most of the 1.67% momentum spread stems from the high performance of past winner watches comprising a highly significantly return of 1.26%. Table 7.5

shows that luxury watches load significantly on market betas. This raises the question of whether portfolio returns are robust after controlling for market exposure. For this reason, Table 7.6 shows (CAPM-) alphas with respect to the value-weighted luxury watch market return. Our results indicate that the momentum return spread of 1.67% p.m. for watches with continuous information is not attributable to market exposure as the related alpha is a highly significant 1.35%.

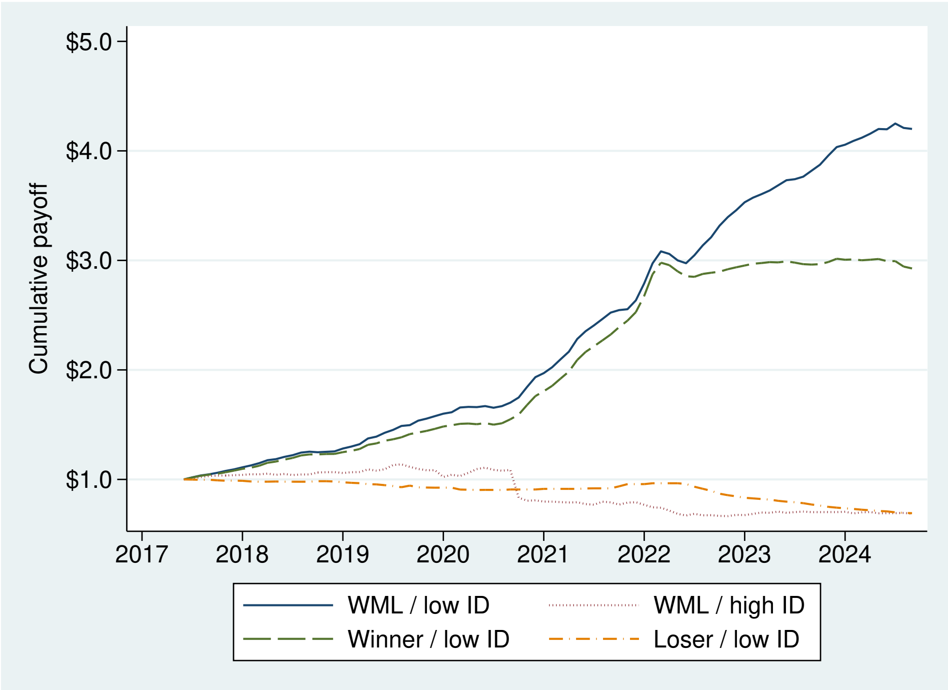

Fig. 7.2. This figure shows the cumulative monthly portfolio returns for selected portfolios from independent, bivariate sorts on momentum MOM and information discreteness ID. The portfolios, which are constructed at the end of each month, are the intersections of three portfolio formed on MOM and three portfolios formed on ID. The breakpoints for both portfolios are the 30% and 70% percentile of MOM and ID, respectively, and we calculate subsequent value-weighted returns. The solid line represents the long-short (WML) returns of top 30% MOM minus low 30% MOM watches while being allocated into the low 30% ID portfolio. The dotted line represents the long-short (WML) returns of top 30% MOM minus low 30% MOM watches while being allocated into the high 30% ID portfolio. The dashed line shows the returns of a long-only investment into top 30% MOM (Winner) and low 30% ID watches. The dash-dotted line shows the returns of a long-only investment into low 30% MOM (Loser) and low 30% ID watches. The sample period is June 2017 to September 2024.

Figure 7.2 visualizes the cumulative profits of selected portfolios formed by our independent, bivariat sorts on MOM and ID. These plots are important because they clearly show that most of the long-short luxury watch momentum strategy returns is generated by the easily investable long leg. The long-short WML strategy among low ID watches (solid line) generates $4.20 until September 2024 for each dollar invested in June 2017 and $2.93 are attributable to the long leg comprising past winner (dashed line). Loser watches that are allocated to the low ID portfolio do not generate any returns until September 2023 and subsequent cumulative losses amount a negligible $0.09 until September 2024 (dashed-dotted line). Clearly, momentum profits are vastly greater among luxury watches with continuous information (i.e., low ID) and most of these profits are generated by the more easily investable long leg of the momentum strategy.

In summary, the results in this section show that *ID* exhibits a negative relation with momentum. Once we control for different levels of information discreteness, the momentum effect is no longer prevalent. Although return consistency adds to the explanatory power of expected luxury watch returns, *ID* seems to be the true effect, for which *RC* is just a proxy. Consistent with the frog-in-the-pan hypothesis which states that investors on the stock market underreact to small amounts of information that arrive continuously (see Da et al. (2014)), we find evidence that this limited attention induced phenomenon is also observed on the market for luxury watches.

7.5 Additional results and further robustness tests

So far, we have investigated the return predictability associated with information discreteness and show that momentum profits can be attributed to *ID*. Our main findings show that luxury watch investors tend to underreact towards continuous information, thus generating a form of long-term mispricing in the sense of Shleifer (1986) and Hirshleifer and Teoh (2003).[7] In this section, we directly test implications of the resulting mispricing and examine market state dependency and the relationship with investor sentiment, because these explanations have also been frequently used in the literature to explain the momentum effect. In addition, we employ the instrumented principal component analysis (IPCA) introduced in Kelly et al. (2019). The method conveniently links conditional alphas and betas with our comprehensive set of luxury watch characteristics and allows to analyze testable implications from our main findings.

7.5.1 Market state and sentiment dependency

Cooper et al. (2004) shows that stock momentum profits depend on the state of the market. If the past three-year cumulative stock market return is negative (down market state), the momentum strategy does not generate significant returns. Vice versa, momentum profits are huge if the three-year cumulative market return is positive (up market state) and Cooper et al. (2004) shows that this finding is likely the effect of investor overconfidence

[7] At least over the four-months momentum formation horizon.

induced mispricing in the meaning of Daniel et al. (1998) and Hong and Stein (1999). Köstlmeier and Röder (2025b) vastly document that mispricing related mechanisms are an important driver of long-short strategy returns, among them momentum. Based on our findings in this study, we hypothesize that luxury watch momentum profits should not depend on market states after controlling for *ID* induced mispricing.

In addition to the market state, Stivers and Sun (2010) show that return dispersion, i.e., the cross-sectional return standard deviation of 10×10 portfolios formed on size and book-to-market, is inversely related with momentum profits (see also Loungani et al. (1990) and Gomes et al. (2003)). The framework presented in Zhang (2005) suggests that return dispersion may contain incremental information about the current state of the economy, beyond market-level returns. Further, if high return dispersion is associated with changes to a weaker market state as suggested in Chordia and Shivakumar (2002) and Cooper et al. (2004), then subsequent momentum payoffs may be lower because past relative performance is unlikely to be sustained with a changing market state. Again, if returns of the luxury watch momentum strategy are, at least partially, associated with mispricing, they should not depend on the level of return dispersion after controlling for *ID*.

Last, long-short returns of many anomalies are stronger following periods of high market-wide investor sentiment (Stambaugh et al. (2012)). As shown in Baker and Wurgler (2006), anomaly returns are affected by time-variation in sentiment if they are associated with market-wide mispricing. Based on our previous findings we hypothesize that luxury watch momentum generates higher returns following periods of high investor sentiment and that return differences are the outcome of *ID* induced mispricing.

Table 7.7 presents our results for the market state, return dispersion, and sentiment dependency of luxury watch momentum portfolio returns. First, Panel A shows results for the momentum strategy *WML*. At the end of each month, luxury watches are allocated to three portfolios based on momentum (*MOM*) using the 30th and 70th percentiles as breakpoints. *WML* denotes the long-short returns which are calculated as the difference between high and low *MOM* value-weighted portfolio returns. To control for information discreteness *ID* in Panel B and C, we use the intersection of 3×3 portfolios based on independent monthly sorts of luxury watches on *MOM* and *ID*. These are the same value-weighted momentum portfolio returns as already presented in Table 7.6. In line with Cooper et al. (2004), the up (down) market state subperiod includes all months t for which the aggregated luxury watch market return over months $t-6$ to $t-1$ is positive (negative).

Table 7.7: Market state and sentiment dependency of luxury watch momentum returns.

This table shows average monthly long-short luxury watch momentum portfolio returns for different market states, return dispersions, and sentiment subperiods. At the end of each month, luxury watches are allocated to three portfolios based on momentum (MOM) using the 30th and 70th percentiles as breakpoints. The long-short returns (WML) are calculated as the difference between high and low value-weighted portfolio returns. To control for ID in our value-weighted momentum portfolios presented in Panel B and Panel C, we use the 3×3 intersection of monthly independent sorts on MOM and ID already presented in Table 7.6. In line with Cooper et al. (2004), the up (down) market state subperiod includes all months t for which the aggregated luxury watch market return over months $t-6$ to $t-1$ is positive (negative). Similar to Stivers and Sun (2010), return dispersion is measured as the cross-sectional return standard deviation of all luxury watch series in our sample. Return dispersion in month t is considered to be high (low), if the average return dispersion of months $t-3$ to $t-1$ is above (below) the time-series median. Following Stambaugh et al. (2012), months t are classified as high (low) sentiment months if the investor sentiment index proposed in Baker and Wurgler (2006) in month $t-1$ is above (below) the time-series median (data retrieved from Jeffrey Wurgler's website). Newey and West (1987) corrected t-statistics using a lag of six are given in parenthesis for subperiod average returns. Δ refers to the difference between the two subperiods and corresponding t-statistics are based on two-sample t-tests (Welchs' t-test with unequal variance). */**/*** indicate significance at the 10%/5%/1% level. The sample period is January 2018 to September 2024.

	Market State		Return Dispersion		Sentiment	
	Up	Down	Low	High	High	Low
Panel A: WML						
mean	1.41***	0.88***	0.84***	1.49***	1.49***	0.85***
	(4.73)	(3.98)	(5.56)	(4.68)	(5.14)	(9.49)
Δ	0.53*		−0.64**		0.65**	
	(1.76)		(−2.01)		(2.13)	
Panel B: WML conditional on low ID						
mean	1.92***	1.18***	1.10***	2.04***	2.04***	1.15***
	(6.34)	(4.08)	(5.16)	(6.73)	(7.47)	(8.98)
Δ	0.74**		−0.94***		0.89***	
	(2.51)		(−3.17)		(3.02)	
Panel C: WML conditional on high ID						
mean	−0.70	0.09	0.14	−0.80	−0.66	0.13
	(−1.44)	(0.53)	(0.91)	(−1.56)	(−1.34)	(1.21)
Δ	−0.79		0.94*		−0.79	
	(−1.50)		(1.68)		(−1.47)	

Similar to Stivers and Sun (2010), return dispersion is measured as the cross-sectional return standard deviation of all luxury watch series in our sample.[8] Return dispersion in month t is considered to be high (low), if the average return dispersion of months $t-3$ to $t-1$ is above (below) the time-series median. Following Stambaugh et al. (2012), months t are classified as high (low) sentiment months if the investor sentiment index proposed in Baker and Wurgler (2006) in month $t-1$ is above (below) the time-series median (data retrieved from Jeffrey Wurgler's website).

Panel A of Table 7.7 shows that the magnitude of the momentum strategy *WML* is much higher after up market states than after down market states. Interestingly, momentum generates highly significantly returns in both market states. This implies that *WML* in the market for luxury watches does not suffer from severe crashes following large market downturns as seen in the stock market (see Daniel and Moskowitz (2016)). Supporting our hypothesis of investor underreaction to continuous information, the difference in *WML* returns between up and down market states increases to a significantly 0.74% among low *ID* watches while being insignificant among high *ID* watches (see Panel B and C).[9]

Beyond the market state, return dispersion is a countercyclical indicator of momentum returns among equities. Table 7.7 shows that momentum profitability depends positively on cross-sectional return dispersion. The difference between periods of low and high return dispersion is a significant -0.64%. While the positive dependency seems surprising at first, it may be the result of our construction of return dispersion relying on the cross-section of all luxury watch returns. As our summary statistics from Table 7.1 indicates, the market return is driven by high returns of few expensive watches. Thus, few high returns of expensive watches may drive the dispersion, and these watches are very likely to enter the momentum portfolio. However, this cannot explain that the difference between high and low return dispersion states is a highly significant -0.94% conditional on low *ID* while being only weekly significant conditional on high *ID*. Our last analysis on investor sentiment shows that *WML* profits are much larger following periods of high sentiment. To put this into perspective, recall that WML generates on average a monthly return

[8] In the original work, the intersection of 10×10 (stock) portfolios based on size and book-to-market are used. Because of data limitations, we are not able to construct a book-to-market ratio for luxury watches. However, using data from K. French data library, our results remain valid even if we use return dispersion based on these stock portfolios.

[9] As robustness tests, we impose short-selling constraints and evaluate only the long leg of the momentum strategy (i.e., past winner) which is more easily replicable by investors. Our results in this section are, however, not affected by this decision. Exemplarily, past winner generate a highly significant 1.72% in up market states and a negligible -0.12% in down market states with the difference of 1.84% being statistically highly significant (t-stat: 7.34).

of 1.25%. Following high sentiment periods, WML yields 1.49% and only 0.85% after low sentiment periods. Considering only low *ID* watches in the construction of WML, we observer a highly significantly return of 2.04% following high sentiment periods. The difference of 0.89% with WML returns following low sentiment periods is also highly significant, in line with the findings in Stambaugh et al. (2012). As expected, momentum among high *ID* watches seems to be independent from sentiment states. Combining our findings on market state and sentiment dependency, the data is in favor or a mispricing based explanation for the momentum effect in the market for luxury watches, where investors underreact to continuous information.

7.5.2 Do momentum and information discreteness matter? An IPCA analysis

IPCA performance

The instrumented principal component analysis (IPCA) introduced in Kelly et al. (2019) offers a novel robustness test for our main findings. To summarize, IPCA represents a fully data-driven method to analyze the common factor structure of given assets. The advantage of IPCA is that it allows its conditional betas to be a linear function of observable characteristics. In the robustness test presented in this section, we use IPCA to retrieve the factor structure in the returns of luxury watch series with time-varying loadings. This allows us to test the following implication from the IPCA framework: If momentum is associated with information discreteness, then both characteristics should enter the same IPCA factor and contribute to this factor to a large extent.[10]

The IPCA model assumes that excess returns $r_{i,t}$ over the period $t = 1, ..., T$ for our N luxury watch series are described by:[11]

$$r_{i,t+1} = \alpha_{i,t} + \beta_{i,t} f_{t+1} + \epsilon_{i,t+1}, \tag{7.2}$$

with f_{t+1} being returns of a total of K latent risk factors. $\epsilon_{i,t+1}$ denotes the residual returns and $\alpha_{i,t}$ and $\beta_{i,t}$ are time-varying intercept (alpha) and conditional factor loadings (betas) for each luxury watch series $i = 1, .., N$. Conditional alphas and betas are given by

[10] In addition, the IPCA analysis also allows us to quantify the overall importance of the momentum characteristic to describe the cross-section of luxury watch returns. So far, we took the relevance of the momentum effect for granted.

[11] We use one-month Treasury Bill rates from K. French data library as a proxy for the risk-free rate.

$$\alpha_{i,t} = z'_{i,t}\Gamma_\alpha + \nu_{\alpha,i,t}, \tag{7.3}$$

$$\beta_{i,t} = z'_{i,t}\Gamma_\beta + \nu_{\beta,i,t}. \tag{7.4}$$

The $L \times 1$ vector Γ_α and the $L \times K$ matrix Γ_β map the total of L luxury watch characteristics summarized in the $L \times 1$ vector $z_{i,t}$ to the mispricing and risk factor exposure, while ν refers to related residuals uncorrelated with watch characteristics. We define $W_t = Z'_t Z_t / N_{t+1}$ to be an $L \times L$ matrix with Z_t representing an $N_{t+1} \times L$ matrix of luxury watch characteristics. To be specific, N_{t+1} denotes the number of non-missing observations of characteristics in month t having a valid return at month $t+1$. For ease of interpretation, if we define an $L \times 1$ vector $X_{t+1} = Z'_t R_{t+1} / N_{t+1}$ with R_{t+1} being a $N_{t+1} \times 1$ vector of luxury watch series excess returns at time $t+1$, than we can interpret X_{t+1} as the $t+1$ returns on a set of L managed portfolios. This only requires that we convert characteristics into ranks prior to parameter estimation. The estimation of the model system using pooled ordinary least squares regression is given by the first order conditions:

$$\hat{f}_{t+1} = \left(\hat{\Gamma}'_\beta W_t \hat{\Gamma}_\beta\right)^{-1} \hat{\Gamma}'_\beta \left(X_{t+1} - W_t \hat{\Gamma}_\alpha\right), \tag{7.5}$$

$$\text{vec}\left(\hat{\Gamma}'\right) = \left(\sum_{t=1}^{T-1} W_t \otimes \hat{\tilde{f}}_{t+1} \hat{\tilde{f}}'_{t+1}\right)^{-1} \sum_{t=1}^{T-1} X_{t+1} \otimes \hat{\tilde{f}}_{t+1}, \tag{7.6}$$

where $\tilde{f}_{t+1} = [f_{t+1} : 1]'$, and $\Gamma = [\Gamma_\beta : \Gamma_\alpha]$.

The pricing performance of the model is calculated by

$$R^2_{\text{total}} = 1 - \frac{\sum_{i,t}\left(r_{i,t+1} - Z'_{i,t}\left(\hat{\Gamma}_\alpha + \hat{\Gamma}_\beta \hat{f}_{t+1}\right)\right)^2}{\sum_{i,t} r^2_{i,t+1}}, \tag{7.7}$$

$$R^2_{\text{predictive}} = 1 - \frac{\sum_{i,t}\left(r_{i,t+1} - Z'_{i,t}\left(\hat{\Gamma}_\alpha + \hat{\Gamma}_\beta \hat{\lambda}\right)\right)^2}{\sum_{i,t} r^2_{i,t+1}}, \tag{7.8}$$

with $\hat{\lambda}$ denoting the time-series mean return of the estimated IPCA factors. While R^2_{total} measures the overall explanatory power of the factor structure, $R^2_{\text{predictive}}$ evaluates the fraction of realized return variation explained by the model's description of conditional expected returns.

We use our 11 variables presented in Table 7.2 as characteristics, but separately consider signed versions of *ID* (i.e., *posID* and *negID*) and *RC* (i.e., *posRC* and *negRC*) as used in Table 7.5 for a more detailed insight into the dynamics between these characteristics and luxury watch series returns. What follows is that we transform all characteristics into cross-sectional ranks from 0 to 1, as is standard in IPCA analysis. We than take a normal inverse cumulative distribution function to convert these ranks into z-scores roughly in the interval $[-2.64, +2.64]$. This is the same approach as presented in Goyal and Saretto (2024) and has the advantage that it increases weights of observations in the tails and thus explicitly considers non-linear relations between extreme characteristics and returns often found in portfolio analysis. In lack of any theoretical or empirical guidance on the number of factors, we perform the IPCA analysis for up to three latent factors. All results presented in this section are based on in-sample estimates of IPCA models because of the limited period from June 2017 to September 2024.

Table 7.8: IPCA model performance.

The table presents performance measures of IPCA models with $K = [1, 2, 3]$ factors for luxury watch returns. Panel A and B report total and predictive R^2 (in %) for the restricted ($\Gamma_\alpha = 0$) and unrestricted ($\Gamma_\alpha \neq 0$) IPCA model. These are calculated with respect to either individual luxury watches (Panel A) or characteristic-managed portfolios (Panel B). Panel C reports p-values, in decimal units, on the bootstrap Wald test of $\Gamma_\alpha = 0$. The sample period is June 2017 to September 2024.

		K		
		1	2	3
Panel A: Individual watches (r_t)				
R^2_{total}	$\Gamma_\alpha = 0$	13.77	20.98	24.60
	$\Gamma_\alpha \neq 0$	18.54	22.87	25.55
$R^2_{\text{predictive}}$	$\Gamma_\alpha = 0$	6.30	1.85	5.17
	$\Gamma_\alpha \neq 0$	5.77	6.24	6.16
Panel B: Managed portfolios (x_t)				
R^2_{total}	$\Gamma_\alpha = 0$	38.06	58.36	67.31
	$\Gamma_\alpha \neq 0$	52.57	63.53	73.23
$R^2_{\text{predictive}}$	$\Gamma_\alpha = 0$	1.28	< 0	4.76
	$\Gamma_\alpha \neq 0$	7.26	10.84	10.89
Panel C: Asset pricing test				
W_α p-value		0.06	0.03	0.30

Table 7.8 reveals that IPCA explains essentially all of the heterogeneity in average luxury watch associated with watch characteristics if at least three factors are included in the specification. With a single factor ($K = 1$), the restricted ($\Gamma_\alpha = 0$) IPCA model explains 13.77% of the total variation in watch returns. We do, however, not have a natural reference point for the performance of the IPCA model because the finance literature lacks an established benchmark model for luxury watches. Nevertheless, the R^2_{total} from the CAPM is 11.9% for the US stock market sample as shown in Kelly et al. (2019). Allowing for $\Gamma_\alpha \neq 0$ increases the R^2_{total} to 18.54%. Overall, our IPCA watch return performance measures compare well with similar estimates for stocks. Our measures are actually higher than those in studies estimating IPCA factors for options (Goyal and Saretto (2024)) or cryptocurrencies (daily, see Bianchi and Babiak (2024)). Among characteristic-managed portfolios x_t, we observe a comparably poor performance even for the $K = 3$ model, indicating that the portfolios are highly non-redundant.

We calculate the Wald statistics $W_\alpha = \hat{\Gamma}'_\alpha \hat{\Gamma}_\alpha$ to evaluate the distance between the restricted ($\Gamma_\alpha = 0$) and unrestricted ($\Gamma_\alpha \neq 0$) model using the bootstrap approach with 1,000 repetitions presented in Kelly et al. (2019). For both a single factor and the two-factor model, the null hypothesis $\Gamma_\alpha = 0$ is rejected at least at the 5% confidence level when testing the joint significance of all characteristics. At $K = 3$, we fail to reject the $\Gamma_\alpha = 0$ hypothesis as indicated by W_α and we notice that the gap in R^2_{total} between the restricted an unrestricted model narrows down to less than one percentage point. The predictive R^2 values are very large and exceed 5.17% for the three-factor model. Recall, that we analyze luxury watches at the series level, thus the behavior of luxury watches aggregated at the series level can differ markedly from individual watches because they may average out a non-negligible fraction of idiosyncratic variation. Overall, our results suggest that a specification with $K = 3$ factors is preferable.

Interpreting IPCA factors

Because the k-th column of Γ_β describes how each characteristic maps into a watch's beta on the k-th factor, according loadings offer interesting insights into the nature of our estimated IPCA factors. Figure 7.3 plots each column of the estimated Γ_β coefficient matrix from the restricted ($\Gamma_\alpha = 0$) IPCA model with $K = 3$ factors.

Loadings on Factor 1 are primarily determined by two characteristics: *MAX* and *IVOL*. These characteristics enter Factor 1's betas with similar magnitudes and opposing signs so that, all else equal, watches with higher maximum returns over the previous month and watches with low idiosyncratic volatility have higher betas on Factor 1. These lottery like

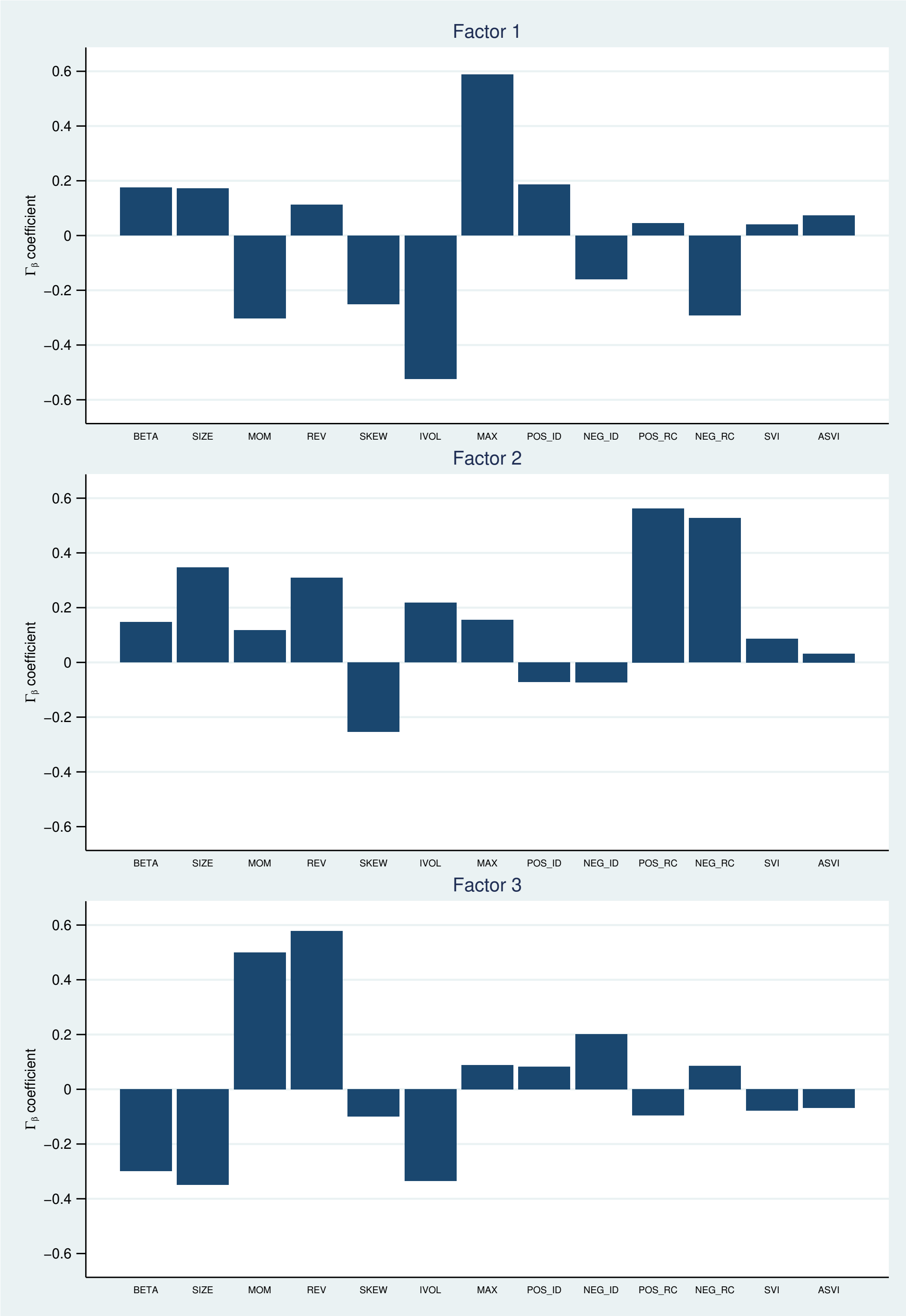

Fig. 7.3. This figure shows each column of the estimated Γ_β coefficient matrix from the restricted ($\Gamma_\alpha = 0$) IPCA model with $K = 3$ factors.

characteristics indicate the prevalence of market wide mispricing among luxury watches, which is in line with previous findings presented in Köstlmeier and Röder (2025b). Factor 1 generates a monthly return of 0.34% p.m. (t-statistics: 4.61) on average. We also notice that Factor 1 negatively loads on momentum and NEG_RC. Taken together, our first factor seems to capture the returns of past loser watches with low idiosyncratic volatility and occasional high returns over the previous month.

Exposure to Factor 2 is mainly driven by return consistency, i.e., both *negRC* and *posRC*. Because momentum enters this factor only moderately, we confirm our previous findings that momentum is unrelated with *RC*. High exposures are obtained for size and reversal which reveals that return consistency is pronounced among expensive watches with high returns over the previous month.

Momentum *MOM* contribution to any IPCA factor is largest for Factor 3. This factor is dominated by momentum and reversal and is further associated with *SIZE*, *BETA*, and *IVOL*. In line with our results, we find that Factor 3 also correlates with *negID*. This means that low degree of information discreteness (i.e., a series of frequent gradual returns) is associate with momentum (and further reversal), whereas contribution of return consistency is much weaker in magnitude for contribution to Factor 3. Despite the relevance of the other characteristics, the analysis in this section, and especially the finding of large contributions of *RC* to Factor 2, shows that return consistency is unlikely to be associated with momentum among luxury watches.

Our IPCA analysis reveals a complex, multivariate dependency of betas on characteristics. First, our results show that return consistency *RC* contributes to a factor (Factor 2) which has only negligible exposure to momentum. This is in line with our previous results that momentum is not likely to be driven by *RC*. Second, watches with a negative exposure to momentum are best captured by Factor 1 which is associated with high *MAX* and low *IVOL*. Third, our estimated Factor 3 represents a combination of momentum and reversal and we find a negative contribution to other characteristics, mainly *SIZE*, *BETA*, and *IVOL*. Notably, exposure to Factor 3 is associated with *negID*. As expected by our analysis in Section 7.4, momentum is associated with information discreteness as both characteristics enter the same IPCA factor. Other than that, the dominance of return consistency and the negligible contribution of momentum in Factor 2 supports our findings that *RC* is very unlikely to drive momentum returns.

7.5.3 Momentum formation period

The main analysis in this study on the momentum effect in the market for luxury watches uses a four-months formation period, skipping the most recent month. This is in line with Köstlmeier and Röder (2025b) who document that this formation period generates a significant return spread using a different sample already beginning in 2010. For robustness of our main results and in line with Jegadeesh and Titman (1993), we consider momentum strategies using past $t \in \{2, ..., 12\}$ to $t-1$ returns while skipping the most recent month in this section.

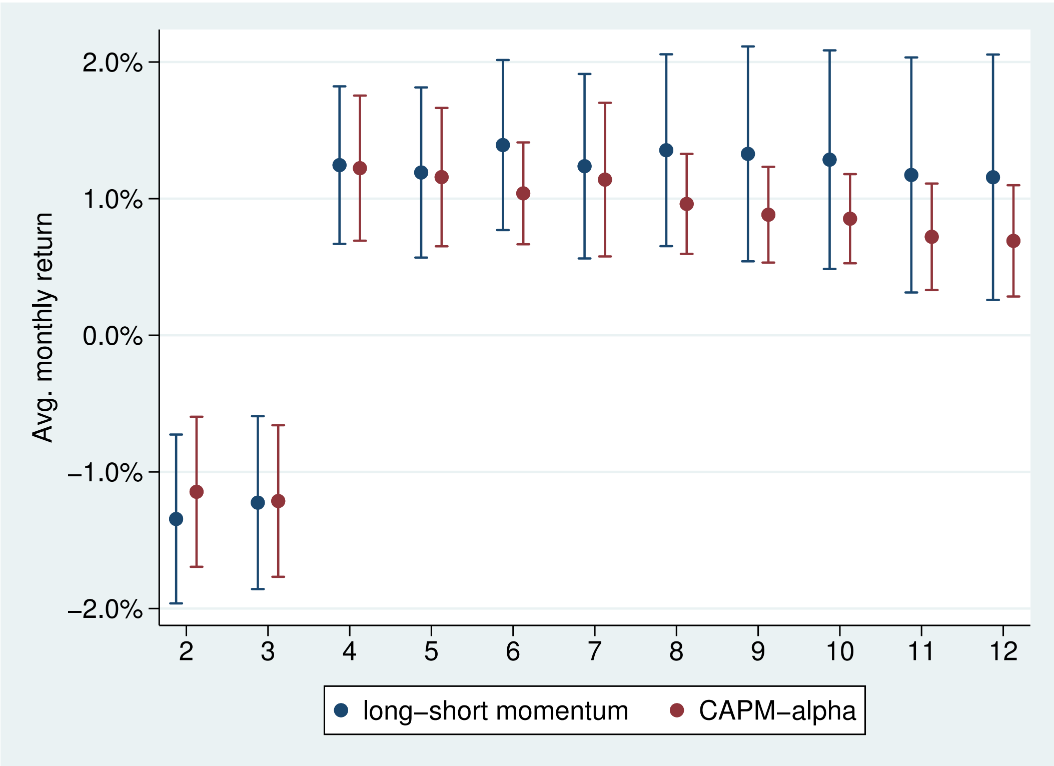

Fig. 7.4. This figure shows average monthly returns of value-weighted long-short momentum portfolio strategies (blue) and related CAPM-alphas (red) with respect to the value-weighted luxury watch market portfolio. The abscissa indicates the momentum portfolio formation months and all strategies skip the most recent month, e.g. column “12” refers to the standard momentum approach following Jegadeesh and Titman (1993) using returns over months $t-12$ to $t-1$. The whiskers correspond to a 99% confidence interval according to Newey and West (1987) robust standard errors. The sample period is June 2017 to September 2024.

Figure 7.4 shows average monthly returns of value-weighted long-short momentum portfolio strategies (blue) and related CAPM-alphas (red) with respect to the value-weighted luxury watch market portfolio. All portfolios are updated each month. The abscissa indicates the momentum portfolio formation months used to calculate MOM.

The whiskers correspond to a 99% confidence interval according to Newey and West (1987) robust standard errors.

We notice that strategies using at least past $t-4$ to $t-1$ returns generate on average a highly significant return of more than one percent per month. The standard momentum approach of Jegadeesh and Titman (1993), i.e., using past $t-12$ to $t-1$ returns generates 1.16% and the related CAPM-alpha is a highly significant 0.69%. Using a very short formation period of up to past three months results in a significant reversal effect, i.e., these strategies generate a highly significant negative return indicating that past loser outperform past winner at this narrow formation period. Overall, our results are robust for the choice of following Köstlmeier and Röder (2025b) and using a past four-months formation period for momentum instead of longer formation periods.

7.5.4 Luxury watch momentum and transaction costs

The costs of trading collectibles may render profits of our luxury watch momentum strategy insignificant. We address this issue in this section by analyzing more feasibly long-only buy/hold strategy implementations and further consider actual costs that occur when selling watches.

To begin with, Figure 7.1 shows that the more easily exploitable long leg contributes the returns to the luxury watch momentum strategy. The high return of 1.15% per month for these watches makes it possible for investors to implement a more feasible long-only approach, forgoing only a negligible 0.10% from the short leg. Novy-Marx and Velikov (2016) demonstrate that trading the momentum anomaly on the stock market exhibits sizable transaction costs. The high turnover of 34.5% per month results in average transaction costs of 0.65%, shrinking the average gross strategy return of 1.33% to 0.68%, which is still achieving statistically significant net returns. To begin our analysis of transaction costs, we take a look on portfolio transitions. We observe that more than 70.1% (74.7%) of watches in the lowest (highest) 30th momentum percentile remain among the lowest (highest) 30th percentile in the subsequent month. The persistence at lag of six months is 51.1% for past winner and even remains at a high level of 46.5% at a lag of 12 months. These results indicate fairly strong persistence of momentum characteristics among luxury watches.

Following Novy-Marx and Velikov (2016), we evaluate an easy-to-implement, rule-based, long-only buy/hold strategy designed to mitigate transaction costs: In the meaning of "sS rules" (see Arrow et al. (1951) and Davis and Norman (1990)), traders buy into a luxury watch when it enters the buy range, but do not sell the watch until it falls out of the hold range, which is larger than the buy range. We implement a conservative 30%/50% buy/hold rule which implies that a trader buys watches when they enter the top 30th percentile of the momentum characteristic, and holds these watches until they fall below the median momentum characteristic.[12] This 30%/50% long-only strategy generates 1.07% p.m. (t-statistic: 3.18) and significantly outperforms the luxury watch market return by a highly significant 0.55%. Many luxury watches are sold on large peer-to-peer platforms like Chrono24.com or WatchCharts.com and buyers may prefer paying through PayPal, since the are eligible for Paypals' Purchase Protection if they pay this way. To account for related fees, we apply a reasonable discount of 2.5% on the price reached by a luxury watch when it is sold according to our 30%/50% buy/hold rule. We observe an average of four sells per month and accounting for a 2.5% transaction fee for each trade still results in a significant momentum strategy return of 0.81% per month, outperforming the luxury watch market return by a significant 0.30%.

7.6 Conclusion

The luxury watch market is a nascent and emerging market. WatchCharts.com, one of the largest peer-to-peer platforms for luxury watches, analyzes millions of data points to determine daily market prices for more than 27,223 luxury watch models from 337 brands. Similar with cryptocurrencies (see Liu et al. (2022)) and other new asset classes that may come into existence in the future, analyzing the market for luxury watches from the empirical asset pricing point of view is important and helps to shed light onto prominent market anomalies.

This paper fills this gap in the literature and studies the market for luxury watches over the period July 2017 to September 2024 through the lens of asset pricing. Using a novel set of data at the daily level, our market index for luxury watches comprises 124 watch series from 26 brands and generates an average monthly return of 0.51% while having a remarkable low volatility of 1.41%. Similar with other asset classes, we document a strong momentum effect on the market for luxury watches generating a highly significantly value-weighted high-low strategy return of 1.25% per month.

This study uses eleven return predictors found to be significant drivers for the cross-section of returns on the stock market and constructs their watch counterparts. Using these characteristics, we offer new insights into the previously documented luxury watch

[12] Novy-Marx and Velikov (2016) suggest a 10%/20% for the stock market momentum strategy.

momentum anomaly (see Köstlmeier and Röder (2025b)) and attribute momentum to limited attention in the meaning of Da et al. (2014): Investors are inattentive to information arriving continuously in small amounts. In other words, a consistent series of frequent gradual returns in luxury watches attracts less attention than occasional large returns and thus giving raise to the momentum effect by inducing strong and persistent return continuation. In line with this frog-in-the-pan hypothesis, we find that momentum returns decrease from a highly significant 1.67% for luxury watches with continuous information during their formation period to an insignificant -0.38% for watches with discrete information, but similar cumulative formation-period returns. Overall, our results suggests that this mispricing induced channel for momentum already documented among stocks (see Da et al. (2014)) is also prevalent in driving momentum strategy returns in the luxury watch market.

Our results have implications for investors, wealth managers, and luxury watch dealers in two ways. First, these people are provided a precise understanding of the risk and returns of the various luxury watch series in which they could invest. Second, the Deloitte Art & Finance Report 2023 states that the wealth of ultra-high-net-worth individuals associated with art and collectibles was already an astonishing \$2.174 trillion in 2022 and is expected to be \$2.861 trillion in 2026, highlighting that an increasing number of people are willing to invest in these alternative investment classes (see Deloitte (2023)). This study may be helpful for related advisors to tilt portfolios of luxury watch investors towards specific investment choices as proxied by our extensive set of analyzed luxury watch characteristics.

Chapter 8

Conclusion

During his time as a managing editor of *The Review of Finance*, A. Edmans rejected more than 1,000 manuscripts and shared his insights from this experience in Edmans (2025). One of his realizations is that submitted research papers often claim to have important implications for policymakers, while often, these implications remain unclear. This is not the case for the studies presented in this dissertation, mainly because our insights on cash flow and discount rate anomaly portfolio shocks in Chapter 2, or on common mispricing in global stock markets based on deviations from related firm-fundamentals in Chapter 3, or on the decomposition of the momentum anomaly into a risk-, mispricing-, and option-component in Chapter 4, have no implications for policymakers. For these types of research papers, A. Edmans emphasizes that it is more important that they change the readers' priors after having read the study. In more detail, *"[. . .] if a reader spends a couple of hours going through your paper, you should change his prior - teach him something new that he did not know or could not have guessed, beforehand."* (p. 421). In this meaning, the main focus of these research papers is to change the readers' prior assumptions about stock market anomalies and stock market mispricing.

In the first research paper, I decompose anomaly portfolio returns using a novel approach proposed in Lochstoer and Tetlock (2020). I extend these previous findings by applying a time-varying vector auto-regression (TVP–VAR) based extended joint connectedness framework as proposed in Balcilar et al. (2021) to estimate spillover effect of related cash flow and discount rate shocks. I find that cash flow shocks explain between 66.64% and 82.65% of anomaly portfolio returns for firms located in the European Monetary Union, and explain 89.71% of market-adjusted return variance at the firm-level, while discount rate shocks are a less important component. In the meaning of Edmans (2025), the readers' prior beliefs I'd like to update by reading this study are the insights that (i) cash flow shocks show a common component among anomaly portfolios, (ii) cash flow shocks show little relation to the business cycle, and (iii), discount rates are the main channel for transmitting spillover effects to other anomalies.

In the second study, I investigate 4,542 firms among 21 global, developed non-U.S. equity markets from June 1990 to June 2021. I begin with extending the fundamentals-based valuation model as proposed in Nichols et al. (2017) to be applicable for a global stock market analysis. The extend version of the model is able to explain 81% of global firms' share price variation. More important, the parsimonious cross-sectional fundamentals-based valuation model links share prices to publicly available accounting fundamentals, so any deviations of observed share prices from the model's derived share prices is helpful to identify mispricing. I document that a related portfolio strategy generates a highly significant 0.56% p.m. and my extensive analysis indicates that this return actually exploits mispricing opportunities rather than being a reward for facing risk exposure.

The third study extends previous insights on stock market mispricing and sets an example based on the momentum effect. In their seminal paper, Jegadeesh and Titman (1993) demonstrate that buying past winner stocks and selling past loser stocks yields positive returns. The robust and large returns generated by momentum strategies are documented among a vast number of asset classes, e.g., mutual funds (Carhart (1997)), commodity futures (Miffre and Rallis (2007)), corporate bonds (Jostova et al. (2013)), cryptocurrencies (Liu et al. (2022)), as well as government bonds and currencies (Asness et al. (2013)), and seems to present a striking contradiction to the weak form of market efficiency hypothesized by Fama (1970). Again, my aim is to update the readers' prior beliefs about the momentum effect, and for that reason, I develop a novel approach for decomposing the returns of 28 U.S. equity momentum strategies into a risk-, mispricing-, and option-component. Average standard momentum returns of 0.66% p.m. contain: (i) an insignificant 0.29% risk-component, (ii) 0.50% mispricing-component, and (iii) -0.13% option-component. The risk-component is related with market volatility and innovations in the term spread, whereas the mispricing-component covaries with illiquidity. While standard factor models capture the risk-component of past losers across all size segments of 5×5 size-momentum portfolios, intercepts for winners are larger in magnitude compared to composed momentum returns.

With regards to the Chapters 5, 6, and 7, it is not very hard to update the readers' prior beliefs, because these studies lay the foundation for the analysis on luxury watches as financial investments.

In the fourth paper, we fill this huge gap in the literature and examine important questions that arise to any investor when being confronted with new investment prospects: Do luxury watches provide additional diversification benefits beyond stocks, bonds, and gold, and if so, are there potential day-of-the-week effects that should be accounted for when buying or selling luxury watches? Using a novel data set comprises daily price data for luxury watches from WatchCharts Analytics from 01/01/2017 to 09/30/2024, we find that some of our analyzed luxury watch indices generate quite large returns. The average annualized return of Rolex (6.94%), Patek Philippe (10.61%), and Audemars Piguet (10.81%) is close to the performance of U.S. stocks (9.28%). We observe that return volatility of luxury watches is remarkably low and just about one fifth of stock market volatility, thus quite the same as Treasury bills. So far, our results clearly show that investors benefit from a diversification potential of luxury watches. Implementing an investment in watches, however, involves to actually buy them, so the question arises: When is the best time, in the sense of which day of the week, to buy them? Highest returns are observed on Wednesdays with 2.42 bps, and returns tend to be generally higher around the mid of a week with average returns of 2.12 bps on Tuesdays and 1.55 bps on Thursdays. Our extensive analysis on these Day-of-the-week-effects further reveals that luxury watch returns are lower on Sundays, because professional dealers typically do not update offers on that day, which is typically a day of rest in most western countries.

In the fifth paper, we move on to analyze a broad sample of 27,289 hand collected watch-month observations from the world's largest peer-to-peer marketplace for luxury watches Chrono24.com between June 2010 and March 2022 through the lens of asset pricing. This unique and novel data set opens new possibilities to test theories of cross-sectional asset pricing anomalies, so we are the first to test 30 characteristics related with the categories size, value, momentum, and volatility in the cross-section of 345 distinct luxury watches from 20 brands. We find that the characteristics size, reversal, short-term momentum, and MAX generate significant difference returns among zero-investment quintile portfolio strategies. Both the k-FWER test method by Lehmann and Romano (2005) and an F-test for the joint significance provide evidence that our results are unlikely to generate by chance. In accordance with the findings of Stambaugh et al. (2015) on the asymmetric pricing effect of sentiment, we find that sentiment-related variation in their performance is mainly due to their short positions. Overall, our results are in favor of a mispricing related interpretation and that the strategies reflect a mispricing commonality across luxury watches.

Last, in the sixth paper, we analyze the luxury watch momentum effect in more detail and provide answers to questions not yet addressed: Why does the momentum effect occur in the market of luxury watches at all? We consider 124 luxury watch indices of 26 brands from WatchCharts.com and their daily returns for the period 06/30/2017 to 09/30/2024. Similar with other asset classes, we document a strong momentum effect generating a highly significant return of 1.25% per month. We find that the inattentiveness of investors to continuously arriving information during the momentum formation period drives momentum returns. This is known as the frog-in-the-pan hypothesis which originates from limited investor attention (see Da et al. (2014)). According to the frog-in-the-pan anecdote, a frog will jump out of a pan containing boiling water since the dramatic temperature change induces an immediate reaction. Conversely, if the water in the pan is slowly raised to a boil, the frog will underreact and perish. Using bivariate independent portfolio sorts, we find that momentum returns decrease from a highly significant 1.67% for luxury watches with continuous information during their formation period to an insignificant -0.38% for watches with discrete information, but similar cumulative formation-period returns. Overall, our battery of empirical tests indicates that indeed limited attention is the core economic channel for the return predictability of continuous information. In conclusion, our results suggests that this mispricing related channel for momentum already documented among stocks (see Da et al. (2014)) is also prevalent in driving momentum strategy returns in the market for luxury watches.

Harvey and Liu (2021) emphasizes that the finance profession has been on a more than 50-year quest to identify factors that explain the cross-section of expected returns. While this refers to the stock market, the research on the cross-section of expected returns in other asset classes, e.g., luxury watches, is still in its infancy. The studies in this dissertation do not provide a definitive truth for all time, but make some progress on a number of vexing questions. The field of asset pricing is developing at a fast pace. New data sets (e.g., high-frequency data) or recent progresses on machine learning techniques used in empirical research promise to bring new insights.

Bibliography

Abalala, T., Sollis, R., 2015. The Saturday effect: an interesting anomaly in the Saudi stock market. *Applied Economics* 47 (58), 6317–6330.

Abdi, F., Ranaldo, A., 2017. A Simple Estimation of Bid-Ask Spreads from Daily Close, High, and Low Prices. *The Review of Financial Studies* 30 (12), 4437–4480.

Abhyankar, A., Filippou, I., Garcia-Ares, P. A., 2024. Overcoming Arbitrage Limits: Option Trading and Momentum Returns. *Journal of Financial and Quantitative Analysis* 59 (1), 97–120.

Aharon, D. Y., Qadan, M., 2019. Bitcoin and the day-of-the-week effect. *Finance Research Letters* 31.

Ahmed, S., Bu, Z., Tsvetanov, D., 2018. Best of the Best: A Comparison of Factor Models. *Journal of Financial and Quantitative Analysis* 54 (4), 1713–1758.

Aït-Sahalia, Y., Parker, J. A., Yogo, M., 2004. Luxury Goods and the Equity Premium. *The Journal of Finance* 59 (6), 2959–3004.

Alquist, R., Israel, R., Moskowitz, T., 2018. Fact, Fiction, and the Size Effect. *The Journal of Portfolio Management* 45 (1), 34–61.

Amihud, Y., 2002. Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets* 5 (1), 31–56.

Analytics, W., 2025. *WatchCharts in Numbers.* [online; October 21, 2025]. URL: https://watchcharts.com/about.

Andersen, T. G., 1996. Return Volatility and Trading Volume: An Information Flow Interpretation of Stochastic Volatility. *The Journal of Finance* 51 (1), 169–204.

Andrei, D., Cujean, J., 2017. Information percolation, momentum and reversal. *Journal of Financial Economics* 123 (3), 617–645.

Andrioli, D. E., May 2020. Will the coronavirus crisis burst the Rolex bubble? *Chrono24 Magazin*. https://www.chrono24.com/magazine/will-the-coronavirus-crisis-burst-the-rolex-bubble-p_63153/ [online; March 28, 2023].

Andrioli, D. E., Feb. 2021. Patek Philippe Ceases Production of the Nautilus 5711: What's Next? *Chrono24 Magazin*. https://www.chrono24.com/magazine/patek-philippe-ceases-production-of-the-nautilus-5711-whats-next-p_74440/ [online; February 18, 2023].

Ang, A., Hodrick, R. J., Xing, Y., Zhang, X., 2006. The Cross-Section of Volatility and Expected Returns. *The Journal of Finance* 61 (1), 259–299.

Ang, A., Hodrick, R. J., Xing, Y., Zhang, X., 2009. High idiosyncratic volatility and low returns: International and further U.S. evidence. *Journal of Financial Economics* 91 (1), 1–23.

Ariel, R. A., 1987. A monthly effect in stock returns. *Journal of Financial Economics* 18 (1), 161–174.

Ariovich, G., 1985. The Economics of Diamond Price Movements. *Managerial and Decision Economics* 6 (4), 234–240.

Arlot, S., Celisse, A., 2010. A survey of cross-validation procedures for model selection. *Statistics Surveys*, 40–79.

Arrow, K. J., Harris, T., Marschak, J., 1951. Optimal Inventory Policy. *Econometrica* 19 (3), 250–272.

Asness, C. S., 2011. Momentum in Japan: The Exception that Proves the Rule. *Journal of Portfolio Management* 37 (4), 67–75.

Asness, C. S., Frazzini, A., 2013. The Devil in HML's Details. *Journal of Portfolio Management* 39 (4), 49–68.

Asness, C. S., Frazzini, A., Israel, R., Moskowitz, T. J., Pedersen, L. H., 2018. Size matters, if you control your junk. *Journal of Financial Economics* 129 (3), 479–509.

Asness, C. S., Frazzini, A., Pedersen, L. H., 2019. Quality minus junk. *Review of Accounting Studies* 24 (1), 34–112.

Asness, C. S., Moskowitz, T. J., Pedersen, L. H., 2013. Value and Momentum Everywhere. *The Journal of Finance* 68 (3), 929–985.

Auer, B. R., Schuhmacher, F., 2013. Diamonds – A precious new asset? *International Review of Financial Analysis* 28, 182–189.

Avery, A. E., Colonna, C. M., 1987. The market for collectible antique and reproduction firearms: An economic and financial analysis. *Journal of Cultural Economics* 11 (2), 49–64.

Avramov, D., Cheng, S., Hameed, A., 2016. Time-Varying Liquidity and Momentum Profits. *Journal of Financial and Quantitative Analysis* 51 (6), 1897–1923.

Avramov, D., Chordia, T., Jostova, G., Philipov, A., 2007. Momentum and Credit Rating. *The Journal of Finance* 62 (5), 2503–2520.

Ayala, M. J., Gonzálvez-Gallego, N., Arteaga-Sánchez, R., 2024. Google search volume index and investor attention in stock market: a systematic review. *Financial Innovation* 10 (1).

Babenko, I., Boguth, O., Tserlukevich, Y., 2016. Idiosyncratic Cash Flows and Systematic Risk. *The Journal of Finance* 71 (1), 425–456.

Bae, K.-H., Wang, W., 2012. What's in a "China" Name? A Test of Investor Attention Hypothesis. *Financial Management* 41 (2), 429–455.

Baker, M., Stein, J. C., 2004. Market liquidity as a sentiment indicator. *Journal of Financial Markets* 7 (3), 271–299.

Baker, M., Wurgler, J., 2006. Investor Sentiment and the Cross-Section of Stock Returns. *The Journal of Finance* 61 (4), 1645–1680.

Baker, M., Wurgler, J., 2007. Investor Sentiment in the Stock Market. *Journal of Economic Perspectives* 21 (2), 129–152.

Baker, M., Wurgler, J., Yuan, Y., 2012. Global, local, and contagious investor sentiment. *Journal of Financial Economics* 104 (2), 272–287.

Balcilar, M., Bouri, E., Gupta, R., Roubaud, D., 2017. Can volume predict Bitcoin returns and volatility? A quantiles-based approach. *Economic Modelling* 64, 74–81.

Balcilar, M., Gabauer, D., Umar, Z., 2021. Crude Oil futures contracts and commodity markets: New evidence from a TVP-VAR extended joint connectedness approach. *Resources Policy* 73, 102219.

Balduzzi, P., Elton, E. J., Green, T. C., 2001. Economic News and Bond Prices: Evidence from the U.S. Treasury Market. *Journal of Financial and Quantitative Analysis* 36 (4), 523–543.

Bali, T. G., Cakici, N., Whitelaw, R. F., 2011. Maxing out: Stocks as lotteries and the cross-section of expected returns. *Journal of Financial Economics* 99 (2), 427–446.

Bali, T. G., Gunaydin, A. D., Jansson, T., Karabulut, Y., 2023. Do the rich gamble in the stock market? Low risk anomalies and wealthy households. *Journal of Financial Economics* 150 (2), 103715.

Ball, R., Gerkos, J., Linnainmaa, J. T., Nikolaev, V. V., 2015. Deflating profitability. *Journal of Financial Economics* 117 (2), 225–248.

Baltussen, G., Swinkels, L., Van Vliet, P., 2021. Global factor premiums. *Journal of Financial Economics* 142 (3), 1128–1154.

Bampinas, G., Fountas, S., Panagiotidis, T., 2016. The day-of-the-week effect is weak: Evidence from the European real estate sector. *Journal of Economics and Finance* 40 (3), 549–567.

Bansal, R., Yaron, A., 2004. Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles. *The Journal of Finance* 59 (4), 1481–1509.

Banz, R. W., 1981. The relationship between return and market value of common stocks. *Journal of Financial Economics* 9 (1), 3–18.

Barber, B. M., Odean, T., 2008. All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors. *Review of Financial Studies* 21 (2), 785–818.

Barberis, N., 2018. *Chapter 2 - Psychology-Based Models of Asset Prices and Trading Volumes*. Vol. 1. Handbook of Behavioral Economics: Applications and Foundations 1. North-Holland, 79–175.

Barberis, N., Jin, L. J., Wang, B., 2021. Prospect Theory and Stock Market Anomalies. *The Journal of Finance* 76 (5), 2639–2687.

Barberis, N., Shleifer, A., 2003. Style investing. *Journal of Financial Economics* 68 (2), 161–199.

Barberis, N., Shleifer, A., Vishny, R., 1998. A model of investor sentiment. *Journal of Financial Economics* 49 (3), 307–343.

Barillas, F., Kan, R., Robotti, C., Shanken, J., 2020. Model Comparison with Sharpe Ratios. *Journal of Financial and Quantitative Analysis* 55 (6), 1840–1874.

Barillas, F., Shanken, J., 2018. Comparing Asset Pricing Models. *The Journal of Finance* 73 (2), 715–754.

Barroso, P., Santa-Clara, P., 2015. Momentum has its moments. *Journal of Financial Economics* 116 (1), 111–120.

Barroso, P., Wang, H., 2025. What Explains Momentum When It Really Works? *Available at SSRN 3716786.*

Barry, C. B., Brown, S. J., 1984. Differential information and the small firm effect. *Journal of Financial Economics* 13 (2), 283–294.

Basu, S., 1977. Investment Performance of Common Stocks in Relation to Their Price-Earnings Ratios: A Test of the Efficient Market Hypothesis. *The Journal of Finance* 32 (3), 663–682.

Bates, D. S., 2019. How Crashes Develop: Intradaily Volatility and Crash Evolution. *The Journal of Finance* 74 (1), 193–238.

Baumol, W. J., 1986. Unnatural Value: Or Art Investment as Floating Crap Game. *The American Economic Review* 76 (2), 10–14.

Baur, D. G., Lucey, B. M., 2010. Is Gold a Hedge or a Safe Haven? An Analysis of Stocks, Bonds and Gold. *Financial Review* 45 (2), 217–229.

Baur, D. G., McDermott, T. K., 2010. Is gold a safe haven? International evidence. *Journal of Banking & Finance* 34 (8), 1886–1898.

Bekaert, G., Harvey, C. R., Lundblad, C., Siegel, S., 2007. Global Growth Opportunities and Market Integration. *The Journal of Finance* 62 (3), 1081–1137.

Belk, R. W., 1988. Possessions and the Extended Self. *Journal of Consumer Research* 15 (2), 139–168.

Beracha, E., Skiba, H., 2011. Momentum in Residential Real Estate. *Journal of Real Estate Finance and Economics* 43 (3), 299–320.

Bergbrant, M. C., Kelly, P. J., 2016. Macroeconomic Expectations and the Size, Value, and Momentum Factors. *Financial Management* 45 (4), 809–844.

Bergmeir, C., Benítez, J. M., 2012. On the use of cross-validation for time series predictor evaluation. *Information Sciences* 191, 192–213.

Berndt, E. R., Savin, E., 1977. Conflict among Criteria for Testing Hypotheses in the Multivariate Linear Regression Model. *Econometrica* 45 (5), 1263–1277.

Bessembinder, H., 2018. Do stocks outperform Treasury bills? *Journal of Financial Economics* 129 (3), 440–457.

Bessembinder, H., Chan, K., Seguin, P. J., 1996. An empirical examination of information, differences of opinion, and trading activity. *Journal of Financial Economics* 40 (1), 105–134.

Bhojraj, S., Swaminathan, B., 2006. Macromomentum: Returns Predictability in International Equity Indices. *The Journal of Business* 79 (1), 429–451.

Bianchi, D., Babiak, M., 2024. Mispricing and Risk Compensation in Cryptocurrency Returns. *Available at SSRN 3935934*.

Birru, J., 2018. Day of the week and the cross-section of returns. *Journal of Financial Economics* 130 (1), 182–214.

Birru, J., Mohrschladt, H., Young, T., 2023. Disentangling Anomalies: Risk versus Mispricing. *Fisher College of Business Working Paper No. 2020-03-029, Charles A. Dice Working Paper No. 2020-29, Available at SSRN 3739944*.

Black, F., 1976. Studies of stock market volatility changes. *Proceedings of the American Statistical Association, Business & Economic Statistics Section, 1976*.

Blitz, D., 2021. The Quant Crisis of 2018-2020: Cornered by Big Growth. *The Journal of Portfolio Management* 47 (6), 8–21.

Blitz, D., Huij, J., Martens, M., 2011. Residual momentum. *Journal of Empirical Finance* 18 (3), 506–521.

Blundell, R., Ward, C. W. R., 1987. Property portfolio allocation: A multi-factor model. *Land Development Studies* 4, 145–156.

Bocart, F. Y. R. P., Hafner, C. M., 2015. Fair Revaluation of Wine as an Investment. *Journal of Wine Economics* 10 (2), 190–203.

Bocart, F. Y. R. P., Hafner, C. M., Kasperskaya, Y., Sagarra, M., 2023. Investing in Superheroes? Comic Art as a New Alternative Investment. *Journal of Alternative Investments* 25 (3), 9–27.

Bogousslavsky, V., Fos, V., Muravyev, D., 2024. Informed Trading Intensity. *The Journal of Finance* 79 (2), 903–948.

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31 (3), 307–327.

Bollerslev, T., Todorov, V., 2023. The jump leverage risk premium. *Journal of Financial Economics* 150 (3), 103723.

Booth, L., Aivazian, V., Demirguc-Kunt, A., Maksimovic, V., 2001. Capital Structures in Developing Countries. *The Journal of Finance* 56 (1), 87–130.

Boston Consulting Group, 2023. *Luxury Preowned Watches, Your Time Has Come.*

Bradshaw, M. T., Richardson, S. A., Sloan, R. G., 2006. The relation between corporate financing activities, analysts' forecasts and stock returns. *Journal of Accounting and Economics* 42 (1), 53–85.

Breusch, T. S., 1979. Conflict Among Criteria for Testing Hypotheses: Extensions and Comments. *Econometrica* 47 (1), 203–207.

Brickley, J. A., 1983. Shareholder wealth, information signaling and the specially designated dividend: An empirical study. *Journal of Financial Economics* 12 (2), 187–209.

Brière, M., Szafarz, A., 2021. When it rains, it pours: Multifactor asset management in good and bad times. *The Journal of Financial Research* 44 (3), 641–669.

Brodeur, A., Gray, D., Islam, A., Bhuiyan, S., 2021. A literature review of the economics of COVID-19. *Journal of Economic Surveys* 35 (4), 1007–1044.

Brooks, R., Del Negro, M., 2005. Country Versus Region Effects in International Stock Returns. *Journal of Portfolio Management* 31 (4), 67–72.

Brunnermeier, M. K., Farhi, E., Koijen, R. S. J., Krishnamurthy, A., Ludvigson, S. C., Lustig, H., Nagel, S., Piazzesi, M., 2021. Review Article: Perspectives on the Future of Asset Pricing. *The Review of Financial Studies* 34 (4), 2126–2160.

Brunnermeier, M. K., Pedersen, L. H., 2009. Market Liquidity and Funding Liquidity. *The Review of Financial Studies* 22 (6), 2201–2238.

Bryzgalova, S., Huang, J., Julliard, C., 2023. Bayesian Solutions for the Factor Zoo: We Just Ran Two Quadrillion Models. *The Journal of Finance* 78 (1), 487–557.

Burnside, C., 2011. The Cross Section of Foreign Currency Risk Premia and Consumption Growth Risk: Comment. *American Economic Review* 101 (7), 3456–3476.

Burton, B. J., Jacobsen, J. P., 1999. Measuring Returns on Investments in Collectibles. *Journal of Economic Perspectives* 13 (4), 193–212.

Butt, H. A., Virk, N. S., 2020. Momentum crashes and variations to market liquidity. *International Journal of Finance & Economics* 27 (2), 1899–1911.

Bybee, L., Kelly, B. T., Manela, A., Xiu, D., 2024. Business News and Business Cycles. *The Journal of Finance* 79 (5), 3105–3147.

Caloia, F. G., Cipollini, A., Muzzioli, S., 2019. How do normalization schemes affect net spillovers? A replication of the Diebold and Yilmaz (2012) study. *Energy Economics* 84, 104536.

Campbell, J., Lo, A., MacKinlay, A. C., Whitelaw, R. F., 1997. *The Econometrics of Financial Markets.* Volume 2. Princeton, NJ: Princeton University Press.

Campbell, J. Y., 1991. A Variance Decomposition for Stock Returns. *The Economic Journal* 101 (405), 157–179.

Campbell, J. Y., 1993. Intertemporal Asset Pricing without Consumption Data. *The American Economic Review* 83 (3), 487–512.

Campbell, J. Y., 1996. Understanding Risk and Return. *Journal of Political Economy* 104 (2), 298–345.

Campbell, J. Y., 2006. Household Finance. *The Journal of Finance* 61 (4), 1553–1604.

Campbell, J. Y., Shiller, R. J., 1988a. Stock Prices, Earnings, and Expected Dividends. *The Journal of Finance* 43 (3), 661–676.

Campbell, J. Y., Shiller, R. J., 1988b. The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors. *The Review of Financial Studies* 1 (3), 195–228.

Campbell, J. Y., Vuolteenaho, T., 2004. Bad Beta, Good Beta. *American Economic Review* 94 (5), 1249–1275.

Campbell, R. A. J., 2008. Art as a Financial Investment. *The Journal of Alternative Investments* 10 (4), 64–81.

Cap Gemini, 2024. *World Wealth Report 2024*. URL: `https://www.worldwealthreport.com`.

Carhart, M. M., 1997. On Persistence in Mutual Fund Performance. *The Journal of Finance* 52 (1), 57–82.

Chaieb, I., Langlois, H., Scaillet, O., 2011. Factors and risk premia in individual international stock returns. *Journal of Financial Economics* 141 (2), 669–692.

Chan, K., Kot, H. W., 2006. Can Contrarian Strategies Improve Momentum Profits. *Journal of Investment Management* 4 (1).

Chan, L. K. C., Jegadeesh, N., Lakonishok, J., 1996. Momentum Strategies. *The Journal of Finance* 51 (5), 1681–1713.

Chen, A. Y., Zimmermann, T., 2022. Open Source Cross-Sectional Asset Pricing. *Critical Finance Review* 11 (2), 207–264.

Chen, N.-F., Roll, R., Ross, S. A., 1986. Economic Forces and the Stock Market. *The Journal of Business* 59 (3), 383–403.

Chen, S.-S., 2012. Revisiting the empirical linkages between stock returns and trading volume. *Journal of Banking & Finance* 36 (6), 1781–1788.

Chen, T., Guestrin, C., 2016. XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794.

Cheon, Y.-H., Lee, K.-H., 2017. Maxing Out Globally: Individualism, Investor Attention, and the Cross Section of Expected Stock Returns. *Management Science* 64 (12), 5807–5831.

Chiarella, C., Kang, B., Nikitopoulos, C. S., Tô, T.-D., 2016. The Return–Volatility Relation in Commodity Futures Markets. *Journal of Futures Markets* 36 (2), 127–152.

Chib, S., 1995. Marginal Likelihood from the Gibbs Output. *Journal of the American Statistical Association* 90 (432), 1313–1321.

Chib, S., 1998. Estimation and Comparison of Multiple Change-Point Models. *Journal of Econometrics* 86 (2), 221–241.

Chib, S., Zeng, X., Zhao, L., 2020. On Comparing Asset Pricing Models. *The Journal of Finance* 75 (1), 551–577.

Chordia, T., Shivakumar, L., 2002. Momentum, Business Cycle, and Time-varying Expected Returns. *The Journal of Finance* 57 (2), 985–1019.

Chrono24, Sept. 2022. The State of the Global Watch Industry. *Chrono24 Magazin*. `https://about.chrono24.com/downloads/press/Chrono24_The_State_of_The_Global_Watch_Industry_Report_No.2.pdf` [online; November 11, 2024].

Chrono24, 2025. *Meet Chrono24*. [online; October 21, 2025]. URL: `https://about.chrono24.com/en/company`.

Chui, A. C. W., Titman, S., Wei, K. C. J., 2010. Individualism and Momentum around the World. *The Journal of Finance* 65 (1), 361–392.

Chung, D. J., 2021. Rolex SA. *Harvard Business School Case*, 521–555.

Clarke, C., 2022. The level, slope, and curve factor model for stocks. *Journal of Financial Economics* 143 (1), 159–187.

Cleveland, W. S., 1979. Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association* 74 (368), 829–836.

Cochrane, J. H., 1991. Production-Based Asset Pricing and the Link Between Stock Returns and Economic Fluctuations. *The Journal of Finance* 46 (1), 209–237.

Cochrane, J. H., 2008. The Dog That Did Not Bark: A Defense of Return Predictability. *The Review of Financial Studies* 21 (4), 1533–1575.

Cochrane, J. H., 2011. Presidential Address: Discount Rates. *The Journal of Finance* 66 (4), 1047–1108.

Cohen, L., Frazzini, A., 2008. Economic Links and Predictable Returns. *The Journal of Finance* 63 (4), 1977–2011.

Cohen, R. B., Polk, C. K., 1998. The impact of industry factors in asset-pricing tests. *Kellogg Graduate School of Management: Working Paper.*

Cohen, R. B., Polk, C. K., Vuolteenaho, T., 2003. The Value Spread. *The Journal of Finance* 58 (2), 609–641.

Cong, L. W., Feng, G., He, J., He, X., 2025. Growing the efficient frontier on panel trees. *Journal of Financial Economics* 167, 104024.

Cong, L. W., Li, Y., Wang, N., 2021. Tokenomics: Dynamic Adoption and Valuation. *The Review of Financial Studies* 34 (3), 1105–1155.

Conrad, J., Kaul, G., 1998. An Anatomy of Trading Strategies. *The Review of Financial Studies* 11 (3), 489–519.

Cooper, I., Mitrache, A., Priestley, R., 2022. A Global Macroeconomic Risk Model for Value, Momentum, and Other Asset Classes. *Journal of Financial and Quantitative Analysis* 57 (1), 1–30.

Cooper, M. J., Gulen, H., Schill, M. J., 2008. Asset Growth and the Cross-Section of Stock Returns. *The Journal of Finance* 63 (4), 1609–1651.

Cooper, M. J., Gutierrez Jr., R. C., Hameed, A., 2004. Market States and Momentum. *The Journal of Finance* 59 (3), 1345–1365.

Corbet, S., Goodell, J. W., Gunay, S., Kaskaloglu, K., 2023. Are DeFi tokens a separate asset class from conventional cryptocurrencies? *Annals of Operations Research* 322 (2), 609–630.

Couts, S. J., Gonçalves, A. S., Rossi, A., 2024. Unsmoothing Returns of Illiquid Funds. *The Review of Financial Studies* 37 (7), 2110–2155.

Da, Z., Engelberg, J., Gao, P., 2011. In Search of Attention. *The Journal of Finance* 66 (5), 1461–1499.

Da, Z., Gurun, U. G., Warachka, M., 2014. Frog in the Pan: Continuous Information and Momentum. *The Review of Financial Studies* 27 (7), 2171–2218.

Da, Z., Liu, Q., Schaumburg, E., 2013. A Closer Look at the Short-Term Return Reversal. *Management Science* 60 (3), 658–674.

Da, Z., Warachka, M. C., 2009. Cashflow risk, systematic earnings revisions, and the cross-section of stock returns. *Journal of Financial Economics* 94 (3), 448–468.

Daniel, K., Hirshleifer, D., Subrahmanyam, A., 1998. Investor Psychology and Security Market Under- and Overreactions. *The Journal of Finance* 53 (6), 1839–1885.

Daniel, K., Hirshleifer, D., Subrahmanyam, A., 2001. Overconfidence, Arbitrage, and Equilibrium Asset Pricing. *The Journal of Finance* 56 (3), 921–965.

Daniel, K., Hirshleifer, D., Sun, L., 2020a. Short- and Long-Horizon Behavioral Factors. *Review of Financial Studies* 33 (4), 1673–1736.

Daniel, K., Moskowitz, T. J., 2016. Momentum crashes. *Journal of Financial Economics* 122 (2), 221–247.

Daniel, K., Mota, L., Rottke, S., Santos, T., 2020b. The Cross-Section of Risk and Returns. *Review of Financial Studies* 33 (5), 1927–1979.

Davis, M. H. A., Norman, A. R., 1990. Portfolio Selection with Transaction Costs. *Mathematics of Operations Research* 15 (4), 676–713.

De Bondt, W. F. M., Thaler, R., 1985. Does the Stock Market Overreact? *The Journal of Finance* 40 (3), 793–805.

De Long, J. B., Shleifer, A., Summers, L. H., Waldmann, R. J., 1990. Noise Trader Risk in Financial Markets. *Journal of Political Economy* 98 (4), 703–738.

DeAngelo, H., DeAngelo, L., 2006. The irrelevance of the MM dividend irrelevance theorem. *Journal of Financial Economics* 79 (2), 293–315.

DeAngelo, H., DeAngelo, L., Stulz, R. M., 2006. Dividend policy and the earned/contributed capital mix: a test of the life-cycle theory. *Journal of Financial Economics* 81 (2), 227–254.

Deloitte, 2023. *Art & Finance Report 2023.*

Deloitte, 2024. *The Deloitte Swiss Watch Industry Study 2023: A calibre of its own.*

DeMiguel, V., Garlappi, L., Uppal, R., 2009. Optimal Versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy? *The Review of Financial Studies* 22 (5), 1915–1953.

den Haan, W. J., Levin, A. T., 2005. Robust Covariance Matrix Estimation with Data-Dependent VAR Prewhitening Order. *NBER Working Paper No. t0255, Available at SSRN 235687.*

Denis, D. J., Osobov, I., 2008. Why do firms pay dividends? International evidence on the determinants of dividend policy. *Journal of Financial Economics* 89 (1), 62–82.

Detzel, A., Novy-Marx, R., Velikov, M., 2023. Model Comparison with Transaction Costs. *The Journal of Finance* 78 (3), 1743–1775.

Diebold, F. X., Yilmaz, K., 2009. Measuring Financial Asset Return and Volatility Spillovers, with Application to Global Equity Markets. *The Economic Journal* 119 (534), 158–171.

Diebold, F. X., Yilmaz, K., 2012. Better to give than to receive: Predictive directional measurement of volatility spillovers. *International Journal of Forecasting* 28 (1), 57–66.

Diebold, F. X., Yilmaz, K., 2014. On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics* 182 (1), 119–134.

Dimson, E., Rousseau, P. L., Spaenjers, C., 2015. The price of wine. *Journal of Financial Economics* 118 (2), 431–449.

Dimson, E., Spaenjers, C., 2011. Ex post: The investment performance of collectible stamps. *Journal of Financial Economics* 100 (2), 443–458.

Dittmar, A. K., 2000. Why Do Firms Repurchase Stock? *The Journal of Business* 73 (3), 331–355.

Dobrynskaya, V., Kishilova, J., 2022. Lego: the toy of smart investors. *Research in International Business and Finance* 59, 101539.

Dorfleitner, G., 2003. Why the return notion matters. *International Journal of Theoretical and Applied Finance* 6 (1), 73–86.

Dorfleitner, G., Lung, C., 2018. Cryptocurrencies from the perspective of euro investors: a re-examination of diversification benefits and a new day-of-the-week effect. *Journal of Asset Management* 19 (7), 472–494.

Dowling, M., 2022. Fertile LAND: Pricing non-fungible tokens. *Finance Research Letters* 44, 102096.

Easterbrook, F. H., 1984. Two Agency-Cost Explanations of Dividends. *The American Economic Review* 74 (4), 650–659.

Eaton, G. W., Green, T. C., Roseman, B. S., Wu, Y., 2022. Retail trader sophistication and stock market quality: Evidence from brokerage outages. *Journal of Financial Economics* 146 (2), 502–528.

Edmans, A., 2025. Learnings From 1000 Rejections. *Financial Management* 54 (2), 419–444.

Ehsani, S., Linnainmaa, J. T., 2022. Factor Momentum and the Momentum Factor. *The Journal of Finance* 77 (3), 1877–1919.

Eichenbaum, M. S., Rebelo, S., Trabandt, M., 2021. The Macroeconomics of Epidemics. *The Review of Financial Studies* 34 (11), 5149–5187.

Elder, B., Jan. 2023. Is that the sound of a luxe watch bubble popping? *Financial Times*. `https://www.ft.com/content/da563f69-6ed3-44a6-a2b1-f56661bbff9b` [online; March 28, 2023].

Engle, R. F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* 50 (4), 987–1007.

Engle, R. F., 2002. Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models. *Journal of Business & Economic Statistics* 20 (3), 339–350.

Engle, R. F., Ng, V. K., 1993. Measuring and Testing the Impact of News on Volatility. *The Journal of Finance* 48 (5), 1749–1778.

Engle, R. F., Sheppard, K., 2001. Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH. *NBER Working Paper*.

Erb, C. B., Harvey, C. R., 2006. The Strategic and Tactical Value of Commodity Futures. *Financial Analysts Journal* 62 (2), 69–97.

Estrada, J., 2016. Alternatives: How? How much? Why? *Journal of Wealth Management* 19 (3), 49–61.

Estrella, A., Hardouvelis, G. A., 1991. The Term Structure as a Predictor of Real Economic Activity. *The Journal of Finance* 46 (2), 555–576.

Eun, C. S., Lee, J., 2010. Mean–variance convergence around the world. *Journal of Banking & Finance* 34 (4), 856–870.

Fama, E. F., 1965. The Behavior of Stock-Market Prices. *Journal of Business* 38 (1), 34–105.

Fama, E. F., 1970. Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance* 25 (2), 383–417.

Fama, E. F., French, K. F., 2004. The Capital Asset Pricing Model: Theory and Evidence. *Journal of Economic Perspectives* 18 (3), 25–46.

Fama, E. F., French, K. R., 1988. Dividend yields and expected stock returns. *Journal of Financial Economics* 22 (1), 3–25.

Fama, E. F., French, K. R., 1989. Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics* 25 (1), 23–49.

Fama, E. F., French, K. R., 1992. The Cross-Section of Expected Stock Returns. *The Journal of Finance* 47 (2), 427–465.

Fama, E. F., French, K. R., 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33 (1), 3–56.

Fama, E. F., French, K. R., 1995. Size and book-to-market factors in earnings and returns. *The Journal of Finance* 50 (1), 131–155.

Fama, E. F., French, K. R., 1996. Multifactor Explanations of Asset Pricing Anomalies. *The Journal of Finance* 51 (1), 55–84.

Fama, E. F., French, K. R., 2007. Migration. *Financial Analysts Journal* 63 (3), 45–58.

Fama, E. F., French, K. R., 2008a. Average returns, B/M, and share issues. *The Journal of Finance* 63 (6), 2971–2995.

Fama, E. F., French, K. R., 2008b. Dissecting Anomalies. *The Journal of Finance* 63 (4), 1653–1678.

Fama, E. F., French, K. R., 2012. Size, Value and Momentum in International Stock Returns. *Journal of Financial Economics* 105 (3), 457–472.

Fama, E. F., French, K. R., 2015. A five-factor asset pricing model. *Journal of Financial Economics* 116 (1), 1–22.

Fama, E. F., French, K. R., 2016. Dissecting Anomalies with a Five-Factor Model. *Review of Financial Studies* 29 (1), 69–103.

Fama, E. F., French, K. R., 2017. International Tests of a Five-Factor Asset Pricing Model. *Journal of Financial Economics* 123 (3), 441–463.

Fama, E. F., French, K. R., 2018. Choosing factors. *Journal of Financial Economics* 128 (2), 234–252.

Fama, E. F., French, K. R., 2020. Comparing Cross-Section and Time-Series Factor Models. *The Review of Financial Studies* 33 (5), 1891–1926.

Fama, E. F., MacBeth, J. D., 1973. Risk, Return, and Equilibrium: Empirical Tests. *Journal of Political Economy* 81 (3), 607–636.

Federation of the Swiss Watch Industry FH, 2024. *World Watchmaking Industry in 2023*.

Feng, G., Giglio, S., Xiu, D., 2020. Taming the Factor Zoo: A Test of New Factors. *The Journal of Finance* 75 (3), 1327–1370.

Ferson Wayne E. Foerster, S. R., Keim, D. B., 1993. General Tests of Latent Variable Models and Mean-Variance Spanning. *The Journal of Finance* 48 (1), 131–156.

Firstenberg, P. M., Ross, S. A., Zisler, R. C., 1988. Real Estate: The Whole Story. *Journal of Portfolio Management* 14 (3), 22–34.

Fisher, R. A., 1915. Frequency Distribution of the Values of the Correlation Coefficient in Samples from an Indefinitely Large Population. *Biometrika* 10 (4), 507–521.

Fisher, T. J., Gallagher, C. M., 2012. New Weighted Portmanteau Statistics for Time Series Goodness of Fit Testing. *Journal of the American Statistical Association* 107 (498), 777–787.

Foster, G., Olsen, C., Shevlin, T., 1984. Earnings Releases, Anomalies, and the Behavior of Security Returns. *The Accounting Review* 59 (4), 574–603.

Frazzini, A., Pedersen, L. H., 2014. Betting against beta. *Journal of Financial Economics* 111 (1), 1–25.

French, K. R., 1980. Stock returns and the weekend effect. *Journal of Financial Economics* 8 (1), 55–69.

Frey, J., 2024. Which stock return predictors reflect mispricing? *Available at SSRN 4584121.*

Fried, J. M., Wang, C. C. Y., 2021. Short-termism, shareholder payouts and investment in the EU. *European Financial Management* 27 (3), 389–413.

Galanti, S., July 2022. Blockchain technology in the luxury watch industry: moving beyond the hype towards effective implementation. *Compact (by KPMG Netherlands)*. `https://kpmg.com/ch/en/home/insights/2022/07/blockchain-in-luxury-watch-industry.html` [online; March 21, 2023].

Gelfand, A. E., Smith, A. F. M., 1990. Sampling-Based Approaches to Calculating Marginal Densities. *Journal of the American Statistical Association* 85 (410), 398–409.

Gelman, A., 2006. Prior distributions for variance parameters in hierarchial models (comment on article by Browne and Draper). *Bayesian Analysis* 1 (3), 515–534.

Geltner, D., 1993. Estimating Market Values from Appraised Values without Assuming an Efficient Market. *The Journal of Real Estate Research* 8 (3), 325–345.

Geman, S., Geman, D., 1984. Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-6 (6), 721–741.

George, T. J., Hwang, C.-Y., 2004. The 52-Week High and Momentum Investing. *The Journal of Finance* 59 (5), 2145–2176.

Gerakos, J., Linnainmaa, J. T., 2018. Decomposing Value. *The Review of Financial Studies* 31 (5), 1825–1854.

Getmansky, M., Lo, A. W., Makarov, I., 2004. An econometric model of serial correlation and illiquidity in hedge fund returns. *Journal of Financial Economics* 74 (3), 529–609.

Gibbons, M. R., Ross, S. A., Shanken, J., 1989. A Test of the Efficiency of a Given Portfolio. *Econometrica* 57 (5), 1121–1152.

Glosten, L. R., Jagannathan, R., Runkle, D. E., 1993. On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *The Journal of Finance* 48 (5), 1779–1801.

Glück, M., Hübel, B., Scholz, H., 2021. Currency Conversion of Fama–French Factors: How and Why. *The Journal of Portfolio Management* 47 (2), 157–175.

Goetzmann, W. N., 1992. The accuracy of real estate indices: Repeat sale estimators. *Journal of Real Estate Finance and Economics* 5, 5–53.

Goetzmann, W. N., 1993. Accounting for Taste: Art and the Financial Markets Over Three Centuries. *The American Economic Review* 83 (5), 1370–1376.

Golder, S. A., Macy, M. W., 2011. Diurnal and Seasonal Mood Vary with Work, Sleep, and Daylength Across Diverse Cultures. *Science* 333 (6051), 1878–1881.

Golubov, A., Konstantinidi, T., 2019. Where Is the Risk in Value? Evidence from a Market-to-Book Decomposition. *The Journal of Finance* 74 (6), 3135–3186.

Gomelsky, V., Jan. 2023. In Watches, 2017 Was a Turning Point. *The New York Times*, 10.

Gomes, J., Kogan, L., Zhang, L., 2003. Equilibrium Cross Section of Returns. *Journal of Political Economy* 111 (4), 693–732.

Gormsen, N. J., Jensen, C. S., 2024. Conditional risk. *Journal of Financial Economics* 162, 103933.

Goyal, A., Jegadeesh, N., Subrahmanyam, A., 2025. Empirical determinants of momentum: a perspective using international data. *Review of Finance* 29 (1), 241–273.

Goyal, A., Saretto, A., 2024. Can Equity Option Returns Be Explained by a Factor Model? IPCA Says Yes. *The Review of Financial Studies* hhae087, 1–38.

Greenwood, R., Laarits, T., Wurgler, J., 2023. Stock Market Stimulus. *The Review of Financial Studies* 36 (10), 4082–4112.

Griffin, J. M., 2002. Are the Fama and French Factors Global or Country Specific? *The Review of Financial Studies* 15 (3), 783–803.

Griffin, J. M., Ji, X., Martin, J. S., 2003. Momentum Investing and Business Cycle Risk: Evidence from Pole to Pole. *The Journal of Finance* 58 (6), 2515–2547.

Griffin, J. M., Kelly, P. J., Nardari, F., 2010. Do Market Efficiency Measures Yield Correct Inferences? A Comparison of Developed and Emerging Markets. *The Review of Financial Studies* 23 (8), 3225–3277.

Grinblatt, M., Han, B., 2005. Prospect theory, mental accounting, and momentum. *Journal of Financial Economics* 78 (2), 311–339.

Grinblatt, M., Moskowitz, T. J., 2004. Predicting stock price movements from past returns: the role of consistency and tax-loss selling. *Journal of Financial Economics* 71 (3), 541–579.

Grullon, G., Roni, M., 2002. Dividends, Share Repurchases, and the Substitution Hypothesis. *The Journal of Finance* 57 (4), 1649–1684.

Grullon, G., Roni, M., 2004. The Information Content of Share Repurchase Programs. *The Journal of Finance* 59 (2), 651–680.

Grundy, B. D., Martin, J. S., 2001. Understanding the Nature of the Risks and the Source of the Rewards to Momentum Investing. *The Review of Financial Studies* 14 (1), 29–78.

Guiso, L., Sapienza, P., Zingales, L., 2008. Trusting the Stock Market. *The Journal of Finance* 63 (6), 2557–2600.

Guo, J., Li, P., Li, Y., 2022. What Can Explain Momentum? Evidence from Decomposition. *Management Science* 68 (8), 6184–6218.

Haddad, V., Kozak, S., Santosh, S., 2020. Factor Timing. *Review of Financial Studies* 33 (5), 1980–2018.

Hahn, J., Lee, H., 2006. Yield Spreads as Alternative Risk Factors for Size and Book-to-Market. *Journal of Financial and Quantitative Analysis* 41 (2), 245–269.

Ham, C. G., Kaplan, Z. R., Leary, M. T., 2020. Do dividends convey information about future earnings? *Journal of Financial Economics* 136 (2), 547–570.

Han, Y., Huang, D., Huang, D., Zhou, G., 2022. Expected return, volume, and mispricing. *Journal of Financial Economics* 143 (3), 1295–1315.

Hanauer, M. X., 2020. A Comparison of Global Factor Models. *Available at SSRN 3546295.*

Hanauer, M. X., Windmüller, S., 2023. Enhanced momentum strategies. *Journal of Banking & Finance* 148, 106712.

Hansen, L. P., 1982. Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica* 50 (4), 1029–1054.

Hanson, S. G., Sunderam, A., 2014. The Growth and Limits of Arbitrage: Evidence from Short Interest. *The Review of Financial Studies* 27 (4), 1238–1286.

Hartigan, J. A., Wong, M. A., 1979. Algorithm AS 136: A K-Means Clustering Algorithm. *Journal of the Royal Statistical Society Series C: Applied Statistics* 28 (1), 100–108.

Harvey, C. R., Liu Yan und Zhu, H., 2016. ... and the Cross-Section of Expected Returns. *The Review of Financial Studies* 29 (1), 5–68.

Harvey, C. R., Liu, Y., 2021. Lucky factors. *Journal of Financial Economics* 141 (2), 413–435.

Harvey, C. R., Siddique, A., 2000. Conditional Skewness in Asset Pricing Tests. *The Journal of Finance* 55 (3), 1263–1295.

Hasbrouck, J., 2009. Trading Costs and Returns for U.S. Equities: Estimating Effective Costs from Daily Data. *The Journal of Finance* 64 (3), 1445–1477.

He, X., Hu, F., 2002. Markov Chain Marginal Bootstrap. *Journal of the American Statistical Association* 97 (459), 783–795.

Heston, S. L., Sadka, R., 2008. Seasonality in the cross-section of stock returns. *Journal of Financial Economics* 87 (2), 418–445.

Hillert, A., Jacobs, H., Müller, S., 2014. Media Makes Momentum. *The Review of Financial Studies* 27 (12), 3467–3501.

Hirshleifer, D., Teoh, S. H., 2003. Limited attention, information disclosure, and financial reporting. *Journal of Accounting and Economics* 36 (1-3), 337–386.

Hoffman, A., Jan. 2023. Rolex Reseller Chrono24 Cuts Jobs as Pre-Owned Watch Prices Fall. *Bloomberg.* `https://www.bloomberg.com/news/articles/2023-01-30/rolex-reseller-chrono24-cuts-jobs-as-pre-owned-watch-prices-fall` [online; March 28, 2023].

Holcblat, B., Lioui, A., Weber, M., 2024. Beyond Arbitrage: Deviations from the Risk-Return Tradeoff. *Chicago Booth Research Paper* (22–11). URL: `https://dx.doi.org/10.2139/ssrn.4064159`.

Hollstein, F., 2022. Local, Regional, or Global Asset Pricing? *Journal of Financial and Quantitative Analysis* 57 (1), 291—320.

Hong, H., Lim, T., Stein, J. C., 2000. Bad News Travels Slowly: Size, Analyst Coverage, and the Profitability of Momentum Strategies. *The Journal of Finance* 55 (1), 265–295.

Hong, H., Stein, J. C., 1999. A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets. *The Journal of Finance* 54 (6), 2143–2184.

Hou, K., Dijk, M. A. van, 2019. Resurrecting the Size Effect: Firm Size, Profitability Shocks, and Expected Stock Returns. *The Review of Financial Studies* 32 (7), 2850–2889.

Hou, K., Mo, H., Xue, C., Zhang, L., 2019. Which Factors? *Review of Finance* 23 (1), 1–35.

Hou, K., Mo, H., Xue, C., Zhang, L., 2021. An Augmented q-Factor Model with Expected Growth. *Review of Finance* 25 (1), 1–41.

Hou, K., Moskowitz, T. J., 2005. Market Frictions, Price Delay, and the Cross-Section of Expected Returns. *The Review of Financial Studies* 18 (3), 981–1020.

Hou, K., Xue, C., Zhang, L., 2015. Digesting Anomalies: An Investment Approach. *The Review of Financial Studies* 28 (3), 650–705.

Hou, K., Xue, C., Zhang, L., 2020. Replicating Anomalies. *The Review of Financial Studies* 33 (5), 2019–2133.

Hu, G. X., Pan, J., Wang, J., 2013. Noise as Information for Illiquidity. *The Journal of Finance* 68 (6), 2341–2382.

Huang, S., 2022. The Momentum Gap and Return Predictability. *The Review of Financial Studies* 35 (7), 3303–3336.

Huber, D., 2022. Liquidity and mispricing. *Available at SSRN 3718411.*

Huberman, G., Kandel, S., 1987. Mean-Variance Spanning. *The Journal of Finance* 42 (4), 873–888.

Hung, M., 2000. Accounting standards and value relevance of financial statements: An international analysis. *Journal of Accounting and Economics* 30 (3), 401–420.

Hung, W., Lin, C.-T., Yang, J. J., 2022. Aggregate 52-week high, limited attention, and time-varying momentum profits. *Journal of Banking & Finance* 141, 106531.

Ince, O. S., Porter, R. B., 2006. Individual equity return data from Thomson Reuters Datastream: Handle with care! *Journal of Financial Research* 29 (4), 463–479.

Israel, R., Moskowitz, T. J., Adrienne, R., Serban, L., 1999. Putting an Academic Factor into Practice: The Case of Momentum. *Working Paper.*

Jacob, M., Jacob, M., 2013. Taxation, Dividends, and Share Repurchases: Taking Evidence Global. *The Journal of Financial and Quantitative Analysis* 48 (4), 1241–1269.

Jacobs, H., 2015. What explains the dynamics of 100 anomalies? *Journal of Banking & Finance* 57, 65–85.

Jacobs, H., Müller, S., 2020. Anomalies across the globe: Once public, no longer existent? *Journal of Financial Economics* 135 (1), 213–230.

Jacobs, H., Müller, S., Weber, M., 2014. How should individual investors diversify? An empirical evaluation of alternative asset allocation policies. *Journal of Financial Markets* 19, 62–85.

Jaffe, J. F., 1989. Gold and gold stocks as investments for institutional portfolios. *Financial Analysts Journal* 45 (2), 53–59.

Jaffe, J. F., Jindra, J., Pedersen, D. J., Voetmann, T., 2020. Can mispricing explain the value premium? *Financial Management* 49 (3), 615–633.

Jagannathan, M., Stephens, C. P., Weisbach, M. S., 2000. Financial flexibility and the choice between dividends and stock repurchases. *Journal of Financial Economics* 57 (3), 355–384.

Jeffreys, H., 1961. *Theory of Probability.* 3rd ed. Oxford: Clarendon Press.

Jegadeesh, N., 1990. Evidence of Predictable Behavior of Security Returns. *The Journal of Finance* 45 (3), 881–898.

Jegadeesh, N., Titman, S., 1993. Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *The Journal of Finance* 48 (1), 65–91.

Jegadeesh, N., Titman, S., 2001. Profitability of Momentum Strategies: An Evaluation of Alternative Explanations. *The Journal of Finance* 56 (2), 699–720.

Jensen, M. C., 1986. Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers. *The American Economic Review* 76 (2), 323–329.

Johnson, T. C., 2002. Rational Momentum Effects. *The Journal of Finance* 57 (2), 585–608.

Joseph, K., Wintoki, M. B., Zhang, Z., 2011. Forecasting abnormal stock returns and trading volume using investor sentiment: Evidence from online search. *International Journal of Forecasting* 27 (4), 1116–1127.

Jostova, G., Nikolova, S., Philipov, A., Stahel, C. W., 2013. Momentum in Corporate Bond Returns. *The Review of Financial Studies* 26 (7), 1649–1693.

Kan, R., Zhou, G., 2012. Tests of Mean-Variance Spanning. *Annals of Economics and Finance* 13 (1), 139–187.

Kanchana, K., 2024. *Trending eBay Watch Listings 2024, V1.* URL: `https://www.kaggle.com/dsv/8012760`.

Kane, A., 1984. Coins: Anatomy of a fad asset. *The Journal of Portfolio Management* 10 (2), 44–51.

Kang, Q., 2024. Mispricing in linear asset pricing models. *Applied Economics*, 1–25.

Kapons, M., Kelly, P., Stoumbos, R., Zambrana, R., 2023. Dividends, trust, and firm value. *Review of Accounting Studies* 28 (3), 1354–1387.

Karolyi, G. A., 2016. Home Bias, an Academic Puzzle. *Review of Finance* 20 (6), 2049–2078.

Karpoff, J. M., 1987. The Relation Between Price Changes and Trading Volume: A Survey. *Journal of Financial and Quantitative Analysis* 22 (1), 109–126.

Kaschny, M., Ruppert, D., Bitzhoefer, A., Doniges, K. A., 2015. Incompatibility Hypothesis: Graphical and Mathematical Explanations. *American Journal of Business, Economics and Management* 3 (4), 177–185.

Kass, R. E., Raftery, A. E., 1995. Bayes Factors. *Journal of the American Statistical Association* 90 (430), 773–795.

Kelly, B. T., Malamud, S., Zhou, K., 2024. The Virtue of Complexity in Return Prediction. *The Journal of Finance* 79 (1), 459–503.

Kelly, B. T., Manela, A., Moreira, A., 2021a. Text Selection. *Journal of Business & Economic Statistics* 39 (4), 859–879.

Kelly, B. T., Moskowitz, T. J., Pruitt, S., 2021b. Understanding momentum and reversal. *Journal of Financial Economics* 140 (3), 726–743.

Kelly, B. T., Pruitt, S., Su, Y., 2019. Characteristics are covariances: A unified model of risk and return. *Journal of Financial Economics* 134 (3), 501–524.

Keloharju, M., Linnainmaa, J. T., Nyberg, P., 2016. Return Seasonalities. *The Journal of Finance* 71 (4), 1557–1590.

Kho, B.-C., 1996. Time-varying risk premia, volatility, and technical trading rule profits: Evidence from foreign currency futures markets. *Journal of Financial Economics* 41 (2), 249–290.

Kim, D., Lee, B.-J., 2023. Shorting costs and profitability of long–short strategies. *Accounting & Finance* 63 (1), 277–316.

Kleine, J., Peschke, T., Wagner, N., 2024. Beyond financial wealth: The experienced utility of collectibles. *The Quarterly Review of Economics and Finance* 97, 101865.

Koenker, R., 2005. *Quantile Regression.* Cambridge: Cambridge University Press.

Koenker, R., Bassett, G. J., 1978. Regression Quantiles. *Econometrica* 46 (1), 33–50.

Koijen, R. S., Moskowitz, T. J., Pedersen, L. H., Vrugt, E. B., 2018. Carry. *Journal of Financial Economics* 127 (2), 197–225.

Kolokolova, O., Le Courtois, O., Xu, X., 2021. Is the index efficient? A worldwide tour with stochastic dominance. *Journal of Financial Markets* 59 (B), 100660.

Konstantinidi, T., 2022. Firm life cycle, expectation errors and future stock returns. *Journal of Banking & Finance* 143, 106591.

Koop, G., Pesaran, M. H., Potter, S. M., 1996. Impulse response analysis in nonlinear multivariate models. *Journal of Econometrics* 74 (1), 119–147.

Korteweg, A., Kräussl, R., Verwijmeren, P., 2016. Does it Pay to Invest in Art? A Selection-Corrected Returns Perspective. *The Review of Financial Studies* 29 (4), 1007–1038.

Köstlmeier, S., Röder, K., 2025a. Diversification benefits of luxury watches and day-of-the-week effects in a seven-day traded market. *Available at SSRN 5235200.*

Köstlmeier, S., Röder, K., 2025b. The Global Market for Luxury Watches and Asset Pricing. *Available at SSRN 4431680.*

Koutmos, D., 2018. Bitcoin returns and transaction activity. *Economics Letters* 167, 81–85.

Kozak, S., Nagel, S., Santosh, S., 2018. Interpreting Factor Models. *The Journal of Finance* 73 (3), 1183–1223.

Krasker, W. S., 1979. The Rate of Return to Storing Wines. *Journal of Political Economy* 87 (6), 1363–1367.

Kristoufek, L., 2015. Power-law correlations in finance-related Google searches, and their cross-correlations with volatility and traded volume: Evidence from the Dow Jones Industrial components. *Physica A: Statistical Mechanics and its Applications* 428, 194–205.

Kuhn, T. S., 2012. *The Structure of Scientific Revolutions*. 50th Edition. Chicago, IL: University of Chicago Press.

Kyung, M., Gill, J., Ghosh, M., Casella, G., 2010. Penalized regression, standard errors, and Bayesian lassos. *Bayesian Analysis* 5 (2), 369–411.

Lai, T.-Y., Wang, K., 1998. Appraisal Smoothing: The Other Side of the Story. *Real Estate Economics* 26 (3), 511–535.

Lakonishok, J., Shleifer, A., Vishny, R. W., 1994. Contrarian Investment, Extrapolation, and Risk. *The Journal of Finance* 49 (5), 1541–1578.

Lakonishok, J., Smidt, S., 1988. Are Seasonal Anomalies Real? A Ninety-Year Perspective. *The Review of Financial Studies* 1 (4), 403–425.

Langelett, G., Wang, Z., 2023. Sealed Collectible Card Game Product as Standalone Investment and Portfolio Diversifier. *Global Journal of Accounting and Finance* 7 (1), 1–25.

Larsen, R. J., Kasimatis, M., 1990. Individual differences in entrainment of mood to the weekly calendar. *Journal of Personality and Social Psychology* 58 (1), 164–171.

Laurs, D., Renneboog, L., 2019. My kingdom for a horse (or a classic car). *Journal of International Financial Markets, Institutions and Money* 58, 184–207.

Lazazzera, M., Sept. 2023. If It's Expensive, Does That Make It a Luxury Watch? *The New York Times*, S2.

Le Fur, E., 2023. Risk and return of classic car market prices: passion or financial investment? *Journal of Asset Management* 24 (1), 59–68.

Le Fur, E., Outreville, J.-F., 2019. Fine wine returns: a review of the literature. *Journal of Asset Management* 20 (3), 196–214.

Lee, C. M. C., Swaminathan, B., 2000. Price Momentum and Trading Volume. *The Journal of Finance* 55 (5), 2017–2069.

Lee, J., Naranjo, A., Sirmans, S., 2021. CDS Momentum: Slow-Moving Credit Ratings and Cross-Market Spillovers. *The Review of Asset Pricing Studies* 11 (2), 352–401.

Lehmann, E. L., Romano, J. P., 2005. Generalizations of the Familywise Error Rate. *The Annals of Statistics* 33 (3), 1138–1154.

Lesmond, D. A., Schill, M. J., Zhou, C., 2004. The illusory nature of momentum profits. *Journal of Financial Economics* 71 (2), 349–380.

Lettau, M., Ludvigson, S. C., 2001. Resurrecting the (C)CAPM: A Cross-Sectional Test When Risk Premia Are Time-Varying. *Journal of Political Economy* 109 (6), 1238–1287.

Lettau, M., Ludvigson, S. C., Wachter, J. A., 2008. The Declining Equity Premium: What Role Does Macroeconomic Risk Play? *The Review of Financial Studies* 21 (4), 1653–1687.

Lewellen, J., 2002. Momentum and Autocorrelation in Stock Returns. *The Review of Financial Studies* 15 (2), 533–564.

Li, W. K., Mak, T. K., 1994. On the squared residual autocorrelations in non-linear time series with conditional heteroskedasticity. *Journal of Time Series Analysis* 15 (6), 627–636.

Lintner, J., 1965. The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *The Review of Economics and Statistics* 47 (1), 13–37.

Liu, L. X., Zhang, L., 2008. Momentum Profits, Factor Pricing, and Macroeconomic Risk. *The Review of Financial Studies* 21 (6), 2417–2448.

Liu, W., Moench, E., 2016. What predicts US recessions? *International Journal of Forecasting* 32 (4), 1138–1150.

Liu, Y., Tsyvinski, A., Wu, X., 2022. Common Risk Factors in Cryptocurrency. *The Journal of Finance* 77 (2), 1133–1177.

Ljung, G. M., Box, G. E. P., 1978. On a measure of lack of fit in time series models. *Biometrika* 65 (2), 297–303.

Lo, A. W., MacKinlay, A. C., 1990. Data-Snooping Biases in Tests of Financial Asset Pricing Models. *The Review of Financial Studies* 3 (3), 431–467.

Locatelli-Biey, M., Zanola, R., 2002. The Sculpture Market: An Adjacent Year Regression Index. *Journal of Cultural Economics* 26 (1), 65–78.

Lochstoer, L. A., Tetlock, P. C., 2020. What Drives Anomaly Returns? *The Journal of Finance* 75 (3), 1417–1455.

Lou, D., Polk, C., 2022. Comomentum: Inferring Arbitrage Activity from Return Correlations. *The Review of Financial Studies* 35 (7), 3272–3302.

Loungani, P., Rush, M., Tave, W., 1990. Stock market dispersion and unemployment. *Journal of Monetary Economics* 25 (3), 367–388.

Lucas, R. E., 1978. Asset Prices in an Exchange Economy. *Econometrica* 46 (6), 1429–1445.

Lucey, B. M., Poti, V., Tully, E., 2006. International portfolio formation, skewness and the role of gold. *Frontiers in Finance and Economics* 3 (1), 49–68.

Ludvigson, S. C., Ma, S., Ng, S., 2021. COVID-19 and the Costs of Deadly Disasters. *AEA Papers and Proceedings* 111, 366–370.

Lundberg, S. M., Lee, S.-I., 2017. A unified approach to interpreting model predictions. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 4768–4777.

MacKinlay, A. C., 1995. Multifactor models do not explain deviations from the CAPM. *Journal of Financial Economics* 38 (1), 3–28.

Maillard, S., 2018. The Ambitious Plan of Chrono24: 12 Disruptions in the Watchmaking Industry. *Europa Star*.

Malloy, C. J., Moskowitz, T. J., Vissing-Jørgensen, A., 2009. Long-Run Stockholder Consumption Risk and Asset Returns. *The Journal of Finance* 64 (6), 2427–2479.

Mandel, B. R., 2009. Art as an Investment and Conspicuous Consumption Good. *American Economic Review* 99 (4), 1653–1663.

Mandelbrot, B., 1963. The Variation of Certain Speculative Prices. *Journal of Business* 36 (4), 394–419.

Markowitz, H., 1952. Portfolio Selection. *The Journal of Finance* 7 (1), 77–91.

Martin, S. G., 2016. The Road Less Traveled: The Case for Collectible Automobiles as an Asset Class. *The Journal of Wealth Management* 19 (3), 131–139.

Maurer, R., Reiner, F., Sebastian, S., 2004. Characteristics of German Real Estate Return Distributions: Evidence from Germany and Comparison to the U.S. and U.K. *Journal of Real Estate Portfolio Management* 10 (1), 59–76.

McCracken, M. W., Ng, S., 2016. FRED-MD: A Monthly Database for Macroeconomic Research. *Journal of Business & Economic Statistics* 34 (4), 574–589.

McLean, R. D., Pontiff, J., 2016. Does Academic Research Destroy Stock Return Predictability? *The Journal of Finance* 71 (1), 5–32.

Mehra, R., Prescott, E. C., 1985. The equity premium: A puzzle. *Journal of Monetary Economics* 15 (2), 145–161.

Menzly, L., Ozbas, O., 2010. Market Segmentation and Cross-predictability of Returns. *The Journal of Finance* 65 (4), 1555–1580.

Merton, R. C., 1973. An Intertemporal Capital Asset Pricing Model. *Econometrica* 41 (5), 867–887.

Michaely, R., Rossi, S., Weber, M., 2021. Signaling safety. *Journal of Financial Economics* 139 (2), 405–427.

Miffre, J., Rallis, G., 2007. Momentum strategies in commodity futures markets. *Journal of Banking & Finance* 31 (6), 1863–1886.

Millenary Watches, 2023. Luxury Watch Prices Dropping – Everything you Need to Know. *Millenary Watches.* `https://millenarywatches.com/luxury-watch-prices-dropping-guide/` [online; March 28, 2023].

Miller, E. M., 1977. Risk, Uncertainty, and Divergence of Opinion. *The Journal of Finance* 32 (4), 1151–1168.

Miller, M. H., Scholes, M. S., 1982. Dividends and Taxes: Some Empirical Evidence. *Journal of Political Economy* 90 (6), 1118–1141.

Misirli, E. U., 2023. Aggregate volatility risk and momentum returns. *European Financial Management*, 1–30.

Moench, E., Stein, T., 2025. Equity premium predictability over the business cycle. *Journal of Financial and Quantitative Analysis*, 1–51.

Morgan Stanley and LuxeConsult, 2024. *Top 50 Swiss Watch Brands for 2024 – Industry Report.*

Moroz, D., Pecchioli, B., 2019. Should You Invest in an Old Bottle of Whisky or in a Bottle of Old Whisky? A Hedonic Analysis of Vintage Single Malt Scotch Whisky Prices. *Journal of Wine Economics* 14 (2), 145–163.

Moskowitz, T. J., 2021. Asset Pricing and Sports Betting. *The Journal of Finance* 76 (6), 3153–3209.

Moskowitz, T. J., Grinblatt, M., 1999. Do Industries Explain Momentum? *The Journal of Finance* 54 (4), 1249–1290.

Mossin, J., 1966. Equilibrium in a Capital Asset Market. *Econometrica* 34 (4), 768–783.

Müller, O. R., 2024. *Morgan Stanley and LuxeConsult publish Swiss watch industry's top 50 companies for 2023.* `https://revolutionwatch.com/morgan-stanley-luxeconsult-2024`. [online; November 11, 2024].

Nagel, S., Xu, Z., 2022. Asset Pricing with Fading Memory. *The Review of Financial Studies* 35 (5), 2190–2245.

Nagel, S., Xu, Z., 2023. Dynamics of subjective risk premia. *Journal of Financial Economics* 150 (2), 103713.

Nelson, D. B., 1991. Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica* 59 (2), 347–370.

Newey, W. K., West, K. D., 1987. A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica* 55 (3), 703–708.

Nguyen, C., Hoang, L., Shim, J., Truong, P., 2020. Internet search intensity, liquidity and returns in emerging markets. *Research in International Business and Finance* 52, 101166.

Nichols, D. C., Wahlen, J. M., Wieland, M. M., 2017. Pricing and Mispricing of Accounting Fundamentals in the Time-Series and in the Cross Section. *Contemporary Accounting Research* 34 (3), 1378–1417.

Nissim, D., Ziv, A., 2001. Dividend Changes and Future Profitability. *The Journal of Finance* 56 (6), 2111–2133.

Novy-Marx, R., 2012. Is momentum really momentum? *Journal of Financial Economics* 103 (3), 429–453.

Novy-Marx, R., 2013. The other side of value: The gross profitability premium. *Journal of Financial Economics* 108 (1), 1–28.

Novy-Marx, R., 2015. How Can a Q-Theoretic Model Price Momentum? *NBER Working Paper No. w20985, Available at SSRN 2572144.*

Novy-Marx, R., Velikov, M., 2016. A Taxonomy of Anomalies and Their Trading Costs. *The Review of Financial Studies* 29 (1), 104–147.

Novy-Marx, R., Velikov, M., 2022. Betting against betting against beta. *Journal of Financial Economics* 143 (1), 80–106.

Nozari, M., 2022. Investment horizon for private-value assets: Evidence from the art market. *The Journal of Financial Research* 45 (2), 229–246.

O'Conner, F. A., Lucey, B. M., Batten, J. A., Baur, D. G., 2015. The financial economics of gold — A survey. *International Review of Financial Analysis* 41, 186–205.

Ohlson, J. A., 2005. On Accounting-Based Valuation Formulae. *Review of Accounting Studies* 10 (2), 323–347.

Pan, W.-F., 2020. Does Investor Sentiment Drive Stock Market Bubbles? Beware of Excessive Optimism! *Journal of Behavioral Finance* 21 (1), 27–41.

Parker, J. A., Julliard, C., 2005. Consumption Risk and the Cross Section of Expected Returns. *Journal of Political Economy* 113 (1), 185–222.

Pastor, L., Stambaugh, R. F., 2003. Liquidity Risk and Expected Stock Returns. *Journal of Political Economy* 11 (3), 642–685.

Patton, A. J., Politis, D. N., White, H., 2009. Correction to "Automatic Block-Length Selection for the Dependent Bootstrap" by D. Politis and H. White. *Econometric Reviews* 28 (4), 372–375.

Patton, A. J., Timmermann, A., 2010. Monotonicity in asset returns: New tests with applications to the term structure, the CAPM, and portfolio sorts. *Journal of Financial Economics* 98 (3), 605–625.

Pearce, S., 1995. *On Collecting: An Investigation into Collecting in the European Tradition.* 1st Edition. London: Routledge.

Peng, L., Xiong, W., 2006. Investor attention, overconfidence and category learning. *Journal of Financial Economics* 80 (3), 563–602.

Pesaran, H. H., Shin, Y., 1998. Generalized impulse response analysis in linear multivariate models. *Economics Letters* 58 (1), 17–29.

Petkova, R., 2006. Do the Fama–French Factors Proxy for Innovations in Predictive Variables? *The Journal of Finance* 61 (2), 581–612.

Piotroski, J. D., 2000. Value Investing: The Use of Historical Financial Statement Information to Separate Winners from Losers. *Journal of Accounting Research* 38, 1–41.

Politis, D. N., Romano, J. P., 1994. The Stationary Bootstrap. *Journal of the American Statistical Association* 89 (428), 1303–1313.

Politis, D. N., White, H., 2004. Automatic Block-Length Selection for the Dependent Bootstrap. *Econometric Reviews* 23 (1), 53–70.

Pompe, J., 1996. An Investment Flash: The Rate of Return for Photographs. *Southern Economic Journal* 63 (2), 488–495.

Pontiff, J., 2006. Costly arbitrage and the myth of idiosyncratic risk. *Journal of Accounting and Economics* 42 (1–2), 35–52.

Porter, M. E., 1980. *Competitive Strategy: Techniques for Analyzing Industries and Competitors*. New York: Free Press.

Qadan, M., Aharon, D. Y., Eichel, R., 2022. Seasonal and Calendar Effects and the Price Efficiency of Cryptocurrencies. *Finance Research Letters* 46, 102354.

Rachev, S., Jašić, T., Stoyanov, S., Fabozzi, F. J., 2007. Momentum strategies based on reward–risk stock selection criteria. *Journal of Banking & Finance* 31 (8), 2325–2346.

Rapach, D. E., Strauss, J. K., Zhou, G., 2013. International Stock Return Predictability: What Is the Role of the United States? *The Journal of Finance* 68 (4), 1633–1662.

Redmond, C. H., Cubbage, F. W., 1988. Portfolio Risk and Returns from Timber Asset Investments. *Land Economics* 64 (4), 325–337.

Reichenbach, F., 2025. Skin in the game: The returns of digital assets from computer games. *Finance Research Letters* 83, 107670.

Renneboog, L., Spaenjers, C., 2012. Hard assets: The returns on rare diamonds and gems. *Finance Research Letters* 9 (4), 220–230.

Rhodes–Kropf, M., Robinson, D. T., Viswanathan, S., 2005. Valuation waves and merger activity: The empirical evidence. *Journal of Financial Economics* 77 (3), 561–603.

Rosenberg, B., Reid, K., Lanstein, R., 1985. Persuasive evidence of market inefficiency. *The Journal of Portfolio Management* 11 (3), 9–16.

Ross, M. H., Zondervan, S., 1989. Capital gains and the rate of return on a Stradivarius. *Economic Inquiry* 27 (3), 529–540.

Rossi, A. S., Rossi, P. E., 1977. Body time and social time: Mood patterns by menstrual cycle phase and day of the week. *Social Science Research* 6 (4), 273–308.

Rouwenhorst, K. G., 1998. International Momentum Strategies. *The Journal of Finance* 53 (1), 267–284.

Rush, R. H., 1968. *Antiques as an Investment*. Prentice-Hall.

Sadka, R., 2006. Momentum and post-earnings-announcement drift anomalies: The role of liquidity risk. *Journal of Financial Economics* 80 (2), 309–349.

Sagi, J. S., Seasholes, M. S., 2007. Firm-specific attributes and the cross-section of momentum. *Journal of Financial Economics* 84 (2), 389–434.

Sakemoto, R., 2025. Time-varying group common factors in the stock market anomalies. *The Financial Review* 60 (2), 481–507.

Schmidt, P. S., Schrimpf, A., Arx, U. von, Wagner, A. F., Ziegler, A., 2017. On the Construction of Common Size, Value and Momentum Factors in International Stock Markets: A Guide with Applications. *Swiss Finance Institute Research Paper, No. 10-58.*

Shanken, J., 1992. On the Estimation of Beta-Pricing Models. *The Review of Financial Studies* 5 (1), 1–33.

Sharpe, W. F., 1964. Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *The Journal of Finance* 19 (3), 425–442.

Shiffrin, R. M., Börner, K., Stigler, S. M., 2018. Scientific progress despite irreproducibility: A seeming paradox. *Proceedings of the National Academy of Sciences* 115 (11), 2632–2639.

Shiller, R. J., 1981. Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends? *American Economic Review* 71 (3), 421–436.

Shiller, R. J., 2014. Speculative Asset Prices. *American Economic Review* 104 (6), 1486–1517.

Shiller, R. J., 2017. Narrative Economics. *American Economic Review* 107 (4), 967–1004.

Shleifer, A., 1986. Do Demand Curves for Stocks Slope Down? *The Journal of Finance* 41 (3), 579–590.

Shleifer, A., Vishny, R. W., 1997. The Limits of Arbitrage. *The Journal of Finance* 52 (1), 35–55.

Smales, L. A., 2021. Investor attention and global market returns during the COVID-19 crisis. *International Review of Financial Analysis* 73, 101616.

Stambaugh, R. F., Yu, J., Yuan, Y., 2012. The short of it: Investor sentiment and anomalies. *Journal of Financial Economics* 104 (2), 288–302.

Stambaugh, R. F., Yu, J., Yuan, Y., 2015. Arbitrage Asymmetry and the Idiosyncratic Volatility Puzzle. *The Journal of Finance* 70 (5), 1903–1948.

Stambaugh, R. F., Yuan, Y., 2017. Mispricing Factors. *Review of Financial Studies* 30 (4), 1270–1315.

Statista, 2022. *Retail market share of the leading Swiss watch brands worldwide in 2022.* [online; February 18, 2025]. URL: `https://www.statista.com/statistics/940579/market-share-of-swiss-watch-brands-worldwide/`.

Stivers, C., Sun, L., 2010. Cross-Sectional Return Dispersion and Time Variation in Value and Momentum Premiums. *Journal of Financial and Quantitative Analysis* 45 (4), 987–1014.

Stone, A. A., Schneider, S., Harter, J. K., 2012. Day-of-week mood patterns in the United States: On the existence of 'Blue Monday', 'Thank God it's Friday' and weekend effects. *The Journal of Positive Psychology* 7 (4), 306–314.

Swamy, V., Dharani, M., 2019. Investor attention using the Google search volume index – impact on stock returns. *Review of Behavioral Finance* 11 (1), 56–70.

Tang, W., Zhu, L., 2017. How security prices respond to a surge in investor attention: Evidence from Google Search of ADRs. *Global Finance Journal* 33, 38–50.

Tauchen, G. E., Pitts, M., 1983. The Price Variability-Volume Relationship on Speculative Markets. *Econometrica* 51 (2), 485–505.

The Royal Swedish Academy of Sciences, 2013. *The Prize in Economic Sciences 2013.* [online; October 15, 2025]. URL: `https://www.nobelprize.org/prizes/economic-sciences/2013/summary/`.

Tibshirani, R., Walther, G., Hastie, T., 2001. Estimating the Number of Clusters in a Data Set Via the Gap Statistic. *Journal of the Royal Statistical Society Series B: Statistical Methodology* 63 (2), 411–423.

Tsai, I.-C., 2014. Ripple effect in house prices and trading volume in the UK housing market: New viewpoint and evidence. *Economic Modelling* 40, 68–75.

van Binsbergen, J. H., Boons, M., Opp, C. C., Tamoni, A., 2023. Dynamic asset (mis)pricing: Build-up versus resolution anomalies. *Journal of Financial Economics* 147 (2), 406–431.

van Dijk, M. A., 2011. Is size dead? A review of the size effect in equity returns. *Journal of Banking & Finance* 35 (12), 3263–3274.

Vayanos, D., Woolley, P., 2013. An Institutional Theory of Momentum and Reversal. *The Review of Financial Studies* 26 (5), 1087–1145.

Vuolteenaho, T., 2002. What Drives Firm-Level Stock Returns? *The Journal of Finance* 57 (1), 233–264.

Walkshäusl, C., 2014. The MAX effect: European evidence. *Journal of Banking & Finance* 42, 1–10.

Walkshäusl, C., 2016. Mispricing and the five-factor model. *Economics Letters* 147, 99–102.

Walkshäusl, C., 2021. Predicting stock returns from the pricing and mispricing of accounting fundamentals. *The Quarterly Review of Economics and Finance* 81, 253–260.

Wang, F., Yan, X. S., 2021. Downside risk and the performance of volatility-managed portfolios. *Journal of Banking & Finance* 131, 106198.

Wang, K. Q., Xu, J., 2015. Market volatility and momentum. *Journal of Empirical Finance* 30, 79–91.

Watkins, B., 2003. Riding the Wave of Sentiment: An Analysis of Return Consistency as a Predictor of Future Returns. *Journal of Behavioral Finance* 4 (4), 191–200.

Welch, I., Goyal, A., 2008. A Comprehensive Look at The Empirical Performance of Equity Premium Prediction. *The Review of Financial Studies* 21 (4), 1455–1508.

White, H., 1980. A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica* 48 (4), 817–838.

White, H., 1982. Maximum Likelihood Estimation of Misspecified Models. *Econometrica* 50 (1), 1–25.

Wiest, T., 2023. Momentum: what do we know 30 years after Jegadeesh and Titman's seminal paper? *Financial Markets and Portfolio Management* 37 (1), 95–114.

Zhang, L., 2005. The Value Premium. *The Journal of Finance* 60 (1), 67–103.

Zhang, X. F., 2006. Information Uncertainty and Stock Returns. *The Journal of Finance* 61 (1), 105–137.