Three Essays in Financial Markets

by Pedro Monteiro e Silva Barroso

A dissertation submitted to

Faculdade de Economia

Universidade Nova de Lisboa

in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

in

Finance

July, 2012

Contents

Introduction **1**

1 Beyond the Carry Trade: Optimal Currency Portfolios **6**

1.1 Introduction 6

1.2 Optimal parametric portfolios of currencies 13

1.3 Empirical analysis 19

1.3.1 Data 20

1.3.2 Pre-sample results 26

1.3.3 Out-of-sample results 29

1.4 Comparison with naive currency strategies 33

1.5 Risk exposures 37

1.6 Value to diversified investors 42

1.7 Speculative capital 45

1.8 Conclusion 47

2 The Bottom-up Beta of Momentum **59**

2.1 Introduction 59

2.2 The time-varying beta of momentum 61

2.3 Exposure to the Fama-French factors 67

2.4 Systematic risk and momentum crashes 69

2.5 Conclusion 71

3 Managing the Risk of Momentum **78**

3.1 Introduction 78

3.2 Momentum in the long run 83

3.3 The time-varying risk of momentum 86

3.4 Risk-managed momentum 89

3.5 Economic Significance: An Investor Perspective 92

3.6 Anatomy of momentum risk 97

3.7 Conclusion 99

3.8 Annex: Data sources 100

Bibliography **112**

List of Tables

1.1 The in-sample performance of the investment strategies in the period 1976:02 to 1996:02. 48

1.2 The statistical significance of the variables in the in-sample period of 1976:02 to 1996:02. 49

1.3 The OOS performance of the investment strategies in the period 1996:03 to 2011:12 50

1.4 The performance of naive portfolios in the OOS period compared to the optimal strategy. 51

1.5 The OOS performance of the optimal strategy regressed on the naive portfolios. 51

1.6 Risk exposures of the optimal strategy. 52

1.7 Time-varying risk of the optimal strategy. 53

1.8 The OOS performance of portfolios combining a currency strategy with different background assets. 53

2.1 Descriptive statistics of different betas of momentum. . . 72
2.2 Performance of time-varying betas in explaining excess returns and risk of the momentum strategy. 72
2.3 Descriptive statistics of the conditional betas for the Fama-French model. 73
2.4 Performance of time-varying betas with respect to the Fama-French factors in explaining excess returns and risk. 76
2.5 Performance of the hedged portfolios. 76

3.1 The long-run performance of momentum compared to the Fama-French factors. 102
3.2 AR (1) of 1-month realized variances. 102
3.3 Momentum and the economic gains from scaling. 103
3.4 The economic performance of momentum for a representative investor. 103
3.5 Risk decomposition of momentum risk. 103

List of Figures

1.1 The performance of Deutsche Bank currency ETFs (in euros). 54

1.2 The estimates of the coefficients of the portfolio in the OOS period from 1996:03 to 2011:12. 55

1.3 OOS performance of the strategies versus naive portfolios in the period of 1996:03 to 2011:12. 56

1.4 The OOS value of currency strategies for investors exposed to different background risks. 57

2.1 Bottom-up and unconditional beta of the WML portfolio. 73

2.2 The bottom-up market beta of the WML portfolio and the previous return on the market portfolio. 74

2.3 The loading of the WML on the FF factors (bottom-up betas). 75

3.1 The long-run cumulative returns of momentum compared to the Fama-French factors. 104
3.2 Momentum crashes. 105
3.3 The long-run performance of risk-managed momentum. . 106
3.4 The realized volatility of momentum. 107
3.5 The density of plain momentum and risk-managed momentum. 108
3.6 The benefits of risk-management in the 1930's and the 00's. 109
3.7 Weights of the scaled momentum. 110

Acknowledgements

The PhD program is a long process. I remember the first day of classes but not as if it were yesterday. Many people helped me along the way and I'm deeply grateful to them.

First, I am thankful to Pedro Santa-Clara, my supervisor. Besides being an inspiring professor, I felt encouraged from the onset to follow on his example and face research work with a bold attitude and an unique combination of rigor, creativity and critical thinking. I learned a great deal on how to do research with him, particularly when I had the chance to become his coauthor, a honour I treasure. In particular, I am grateful as he went beyond the role of a supervisor to provide crucial support during some rough patches. He was completely available and always insightful.

I am grateful to Ana Balcão Reis, who guided me through the taught part of the program as PhD coordinator. She was always supportive and helped me find solutions to the most various problems that emerged. All professors of the taught part were always available to help me progress as a student. I would specially mention Iliyan Georgiev. The learning experience in his courses is unforgettable. I am also grateful to Paula Fontoura, the supervisor of my master's dissertation, who first encouraged me to do research work.

I am grateful to colleagues: Arash Aloosh, Daniel Monteiro, Luís Fitchner, José Faias, Rúben Branco and Tymur Gabuniya. Research work is lonely without someone to share experiences. Specially Arash and Daniel, doing the group work with them as students in Pedro Santa-Clara's course was a great and life changing experience for me. José Faias comments and advises were precious.

Financial support from FCT (doctoral grant SFRH/BD/42087/2007) is kindly acknowledged.

Finally, and most of all, I am grateful to my family. My mother Ana for her example of perseverance in the face of adversity, my father Pedro for teaching me most of the values I hold dear, my brother Nuno for the joy he brings to my life and my dear wife Catarina who makes me feel lucky everyday. I dedicate my dissertation to them and the memory of my grandmother Flora Filipa.

Introduction

My PhD thesis consists of three independent essays: one on extensions to the carry trade and two on stock momentum. As independent essays they do not share much in common, but there is an underlying theme: financial market anomalies. In general anomalies are interesting because they should not exist. As such they are of obvious relevance in any scientific field.

In the case of financial markets, returns should be a reward for risk. Otherwise there would be no investors willing to bear those risks. But some strategies seem to offer returns with only elusive relations with fundamental sources of risk.

For example, the carry trade borrows from low interest rate currencies to invest in high yielding ones. Standard macroeconomics manuals explain that interest rate differentials should forecast offsetting movements in currencies. Empirically that is not the case. Hence the returns of the carry trade seem anomalous.

Investment strategies exploiting the information in past returns should not produce abnormal profits either. That seems to violate weak-form market efficiency. But exploiting past information in returns is exactly what momentum does.

Of course, what seems anomalous is not necessarily so. The returns of carry or stock momentum can be rewards for running some misunderstood set of risks.

For instance, both the carry trade and stock momentum exhibit large crash risk. These strategies can be characterized as "picking up pennies in front of a steamroller". A recurrent theme in my thesis is whether the threat of the steamroller explains the persistency of these anomalies.

My first essay is on the carry trade. It is a coauthored chapter with my supervisor, Pedro Santa-Clara. Most research on the carry trade focuses on that strategy alone, improving the understanding of both its risk and returns.

The intuition behind our approach is that there is no obvious reason investors in currencies should restrict themselves to carry strategies. There is pervasive evidence of value and momentum effects across asset classes and we test its relevance in currency investments. It turns out that these contribute a lot to portfolio performance.

The resulting optimal portfolio outperforms the carry trade and other naive benchmarks in out-of-sample tests. Its returns are not explained by risk and are valuable to diversified investors holding stocks and bonds. One important result is that crash-risk is a poor explanation for these returns. A currency speculator can combine the carry with other approaches, obtaining less crash risk. Most notably, a portfolio combining stocks and bonds with diversified currency investments has *less* crash risk than one without currencies.

The second essay is on the time-varying beta of the momentum strategy. Previous research shows some convincing evidence that the beta of momentum changes over time. It was puzzling to me why this previous research did not estimate the betas the most straightforward way: bottom-up from the betas of individual stocks in the portfolio. I do just that and find the unconditional beta of momentum is highly misleading. The bottom-up beta of momentum, estimated from the betas of individual stocks, varies substantially over time. Using bottom-up betas explains up to 40% of the risk of momentum, out-of-sample. This is 17 times more than one unconditional model achieves and outperforms other measures of time-varying beta. But like previous research, I find that hedging in real time the time-varying systematic risk of momentum does not avoid its crashes.

The third essay is on the time-varying volatility of momentum. As the first essay, this is a coauthored chapter with my supervisor. In a first step we wanted to assess if an investor with reasonable risk aversion finds momentum attractive in spite of its rare but intense crashes. We found the answer to be negative. But then we noticed the crashes of momentum are not at all like those in the overall stock market. Momentum risk is quite predictable. The major source of predictability does not come from systematic risk but from momentum-specific risk. Managing this time-varying risk virtually eliminates crashes and nearly doubles the Sharpe ratio of the strategy. We argue that momentum per se is not a very interesting strategy. It provides long runs of attractive returns. But it has also produced sudden crashes that took 30 years or more to recover from. By contrast, risk-managed momentum is more of a puzzle.

Overall, the results of my research do not support the steamroller hypothesis, either for the carry or the stock momentum strategy.

The carry trade crashes. But in a mean-variance world we know risk is not about variance but rather about covariance. Similarly, in a world of fat tails and crashes, risk should not be about the crashes per se but rather about "co-crashes". As a matter of fact the carry does not crash simultaneously with value and momentum. Quite the opposite.

In the case of stock momentum, the steamroller does speed up some times, but it seems to honk in advance.

Chapter 1

Beyond the Carry Trade: Optimal Currency Portfolios[0]

1.1 Introduction

Currency spot rates are nearly unpredictable out of sample (Meese and Rogoff (1983)).[1] Usually, unpredictability is seen as evidence supporting market efficiency, but with currency spot rates it is quite the opposite – it presents a challenge. Since currencies have different interest rates, if the difference in interest rates does not forecast an offsetting depreciation, then investors can borrow the low yielding currencies to invest in the high

[0]We thank Andreas Schrimpf, Adrien Verdelhan, Geert Bekaert, Hanno Lustig, Harald Hau, Maik Schmeling and Nick Roussanov as well as seminar participants at NOVA, the 2011 QED meeting, the 2012 European Winter Finance Summit in Davos, the 2012 INFER annual meeting, and three anonymous referees for their helpful comments and suggestions. We also thank Craig Burnside and Matti Suominen for letting us use their data.

[1]See also Cheung, Chinn, and Pascual (2005), Rogoff and Stavrakeva (2008), Rogoff (2009).

yielding ones (Fama (1984)). This strategy, known as the carry trade, has performed extremely well and for a long period without any sensible economic explanation. Burnside, Eichenbaum, and Rebelo (2008) show that a well-diversified carry trade attains a Sharpe ratio that is more than double that of the US stock market – itself a famous puzzle (Mehra and Prescott (1985)).

Considerable effort has been devoted to explaining the returns of the carry trade as compensation for risk. Lustig, Roussanov, and Verdelhan (2011a) show that the risk of carry trades across currency pairs is not completely diversifiable, so there is a systematic risk component. They form an empirically motivated risk factor – the return of high-yielding currencies minus low-yielding currencies (HML_{FX}) – close in spirit to the stock market factors of Fama and French (1992) and show that it explains the carry premium. But the HML_{FX} is itself a currency strategy, so linking its returns to more fundamental risk sources is an important challenge for research in the currency market.

Some risks of the carry trade are well known. High yielding currencies are known to "go up by the stairs and down by the elevator," implying that the carry trade has substantial crash risk. Carry performs worse when there are liquidity squeezes (Brunnermeier, Nagel, and Pederson (2008)) and increases in foreign exchange volatility (Menkhoff, Sarno, Schmeling, and Schrimpf (2011a)). Its risk exposures are also time-varying, increasing in times of greater uncertainty (Christiansen, Ranaldo, and Söderllind (2010)).

Another possible explanation of the carry premium is that there is some "peso problem" with the carry trade – the negative event that justifies its returns may simply have not occurred yet.[2] Using options to hedge away the "peso risk" reduces abnormal returns, lending some support to this view, but the remaining returns depend crucially on the option strategy used for hedging (Jurek (2009)).

Despite our improved understanding of the risk of the carry trade, the fact remains that conventional risk factors from the stock market (market, value, size, momentum) or consumption growth models, do not explain its returns.[3] Indeed, an investor looking for significant abnormal returns with respect to, say, the Fama-French factors (1992), would do very well by just dropping all equities from the portfolio and investing entirely in a passively managed currency carry portfolio instead.

But there is more to the currency market than just the carry trade. Market practitioners follow other strategies, including value and momentum (Levich and Pojarliev (2011)). The benefits of combining these

[2]Barro (2006), Fahri and Gabaix (2007), Gourio, Siemer, and Verdelhan (2011), Burnside, Eichenbaum, Kleshchelski, and Rebelo (2011).

[3]Burnside, Eichenbaum, and Rebelo (2011), Jordà and Taylor (2011).

different approaches became apparent during the height of the financial crisis when events in the currency market assumed historical proportions.[4] Figure 1.1 shows the performance of three popular Deutsche Bank ETFs that track these strategies with the currencies of the G10. From August 2008 to January 2009, the carry ETF experienced a severe crash of 32.6%, alongside the stock market, commodities and high yield bonds. Even so, this crash was not the "peso event" needed to rationalize its previous returns.[5] But in the same period, the momentum ETF delivered a 29.4% return and the value ETF a 17.8% return. So while the carry trade crashed, a diversified currency strategy fared quite well in this turbulent period.

Coincidently, the literature on alternative currency investments saw major developments since 2008. Menkhoff, Sarno, Schmeling and Schrimpf (2011b) document the properties of currency momentum, Burnside (2011) examines a combination of carry and momentum, Asness, Moskowitz, and Pederson (2009) study a combination of value and momentum in currencies (and other asset classes), and Jordà and Taylor (2009) combine carry, momentum and the real exchange rate.

Most of the studies on alternative currency strategies focus on simple,

[4] Melvin and Taylor (2009) provide a vivid narrative of the major events in the currency market during the crisis.

[5] Burnside, Eichenbaum, Kleshchelski, and Rebelo (2011).

equal-weighted portfolios. The choice of simple portfolios is understandable as there is substantial evidence indicating these typically outperform out-of-sample more complex optimized portfolios.[6] However, we find that using the historical data up to 2007, an investor would have no reason to want to equal-weight momentum, value and carry. Optimized portfolios are a closer reflection of the uncertainties faced by investors in real time. Namely, they have to deal with the choice of what signals to use, how to weigh each signal, and how to address measurement error and transaction costs.

To study the risk and return of currency strategies in a more realistic setting, we use the parametric portfolio policies approach of Brandt, Santa-Clara, and Valkanov (2009) and test the relevance of different variables in forming currency portfolios.

First, we use a pre-sample test to study which characteristics matter for investment purposes. We test the relevance of the interest rate spread (and its sign), momentum and three proxies for value: reversal, the real exchange rate, and the current account. Including all characteristics simultaneously in the test, allows us to see which are relevant and which are subsumed by others. Then we conduct a comprehensive out-of-sample (OOS) exercise with 16 years of monthly returns. This

[6] DeMiguel, Garlappi, and Uppal (2009), Jacobs, Müller, and Weber (2010).

aims to minimize forward-looking bias – though it does not eliminate it completely.[7]

We find that the interest rate spread, momentum and reversal create economic value for investors whereas fundamentals as the current account and the real exchange rate don't. The strategy combining the relevant signals increases the Sharpe ratio relative to an equal-weighted carry portfolio from 0.57 to 0.86, out-of-sample and after transaction costs. This is a 0.29 gain, about the same as the Sharpe ratio of the stock market in the same period.

Transaction costs matter in currency markets. Taking transaction costs into account in the optimization further increases the Sharpe ratio to 1.06, a total gain of 0.49 over the equal-weighted carry benchmark. The gains in certainty equivalent are even more expressive as the optimal diversified strategy substantially reduces crash risk.

Unlike the typical result in OOS tests of optimized equity portfolios, we find that the optimized portfolio outperforms all naive benchmarks.[8] Also, the risk factors recently proposed to explain carry returns do not explain the returns of the optimized portfolio, which has monthly α's

[7] After all, would we be conducting the same out-of-sample exercise in the first place if there were no indications in the literature that momentum and value worked in recent years? Still, unlike naive portfolios, our strategy will not invest in these signals more than justified by the historical data up to each moment in time.

[8] Brandt, Santa-Clara, and Valkanov (2009) optimized portfolio of stocks also outperforms OOS naive benchmarks.

ranging between 1.73 and 2.38 percent. So, while these risk factors may have some success explaining carry returns, they struggle to make sense of our optimal currency strategy.

We assess the benefits of diversification across currency investment strategies for investors already exposed to other asset classes. We find an average increase in the Sharpe ratio of 0.51, a much more impressive gain than the one documented in Kroencke, Schindler, and Schrimpf (2011). Furthermore, including the currency strategies in the portfolio consistently reduces fat tails and left skewness. This contradicts crash-risk explanations for returns in the currency market.

Finally, we regress the returns of the optimal strategy on the level of speculative capital in the market, following Jylhä and Suominen (2011). We find evidence that the returns of the strategy decline as the amount of hedge fund capital increases. This suggests that the returns we document constitute an anomaly that is gradually being arbitraged away by hedge funds.

This chapter is structured as follows. In section 1.2 we explain the implementation of parametric portfolios of currencies. Section 1.3 presents the empirical analysis. Section 1.3.1 describes the data and the variables used in the optimization. Sections 1.3.2 and 1.3.3 present the investment performance of the optimal portfolios in and out of sample, respectively. Section 1.4 compares the performance of the optimal portfolio with naive benchmarks. In Section 1.5 we test the risk exposures of the optimal portfolio. In Section 1.6 we assess the value of currency strategies for investors holding stocks and bonds. Section 1.7 discusses possible explanations for the abnormal returns of the strategy, including insufficient speculative capital early in the sample.

1.2 Optimal parametric portfolios of currencies

We optimize currency portfolios from the perspective of an US investor in the forward exchange market. In the forward exchange market, the investor can agree at time t to buy currency i at time $t+1$ for $1/F^i_{t,t+1}$ where $F^i_{t,t+1}$ is the price of one USD expressed in foreign currency units (FCU). Then at time $t+1$ the investor liquidates the position selling the currency for $1/S^i_{t+1}$, where S^i_{t+1} is the spot price of one USD in FCU. The return (in US dollars) of a long position in currency i in month t is:

$$r^i_{t+1} = \frac{F^i_{t,t+1}}{S^i_{t+1}} - 1 \tag{1.1}$$

This is a zero-investment strategy as it consists of positions in the forward market only.[9] We use one-month forwards throughout as is

[9] In reality investors need to post collateral to take positions in forward markets.

standard in the literature.[10] Therefore all returns are monthly and there are no inherited positions from month to month. This also avoids path-dependency when we include transaction costs in the analysis.

We optimize the currency strategies using the parametric portfolio policies approach of Brandt, Santa-Clara, and Valkanov (2009). This method models the weights of assets as a function of their characteristics. The implicit assumption is that the characteristics convey all relevant information about the assets' conditional distribution of returns. The weight on currency i at time t is:

$$w_{i,t} = \theta^T x_{i,t}/N_t \tag{1.2}$$

where $x_{i,t}$ is a $k \times 1$ vector of currency characteristics, θ is a $k \times 1$ parameter vector to be estimated and N_t is the number of currencies available in the dataset at time t. Dividing by N_t keeps the policy stationary (see Brandt, Santa-Clara, and Valkanov (2009)). We do not place any restriction on the weights, which can be positive or negative. This reflects the fact that in the forward exchange market there is no obvious non-negativity constraint.

The strategies we examine consist of an investment of 100% in the

We ignore that in this study.

[10]Burnside, Eichenbaum, and Rebelo (2008), Burnside (2011), Burnside, Eichenbaum, Kleshchelski, and Rebelo (2011), Menkhoff, Sarno, Schmeling and Schrimpf (2011a,b).

US risk-free asset, yielding rf_t^{US}, and a long-short portfolio in the forward exchange market. For a given sample, θ uniquely determines a parametric portfolio policy, and the corresponding return each period will be:

$$r_{p,t+1} = rf_t^{US} + \sum_{i=1}^{N_t} w_{i,t} r_{t+1}^{i} \tag{1.3}$$

The problem an investor faces is optimizing its objective function picking the best possible θ for the sample:

$$\max_{\theta} E_t \left[U(r_{p,t+1}) \right] \tag{1.4}$$

We use power utility as the objective function:

$$U(r_p) = \frac{(1 + r_p)^{1-\gamma}}{1 - \gamma} \tag{1.5}$$

where γ is the coefficient of relative risk aversion (CRRA).[11] The main advantage of this utility function is that it penalizes kurtosis and skewness, as opposed to mean-variance utility, which focuses only on the first two moments of the distribution of returns. So our investor dislikes crash risk and values characteristics that help reduce it, even if these do not

[11] Bliss and Panigirtzoglou (2004) estimate γ empirically from risk-aversion implicit in one-month options on the S&P and the FTSE and find a value very close to 4. We adopt this value and keep it throughout. The most important measures of economic performance of the strategy are scale-invariant (Sharpe ratio, skewness, kurtosis), so the specific choice of CRRA utility is not of crucial importance.

add to the Sharpe ratio.

The main restriction imposed on the investor's problem is that θ is kept constant across time. This substantially reduces the chances for in-sample overfitting as only a $k \times 1$ vector of characteristics is estimated. The assumption that θ does not change allows its estimation using the sample counterparts:

$$\hat{\theta} = \arg\max_{\theta} \frac{1}{T} \sum_{t=0}^{T-1} U \left(rf_t^{US} + \sum_{i=1}^{N_t} (\theta^T x_{i,t}/N_t) r_{t+1}^i \right) \tag{1.6}$$

For statistical inference purposes, Brandt, Santa-Clara, and Valkanov (2009) show that we can use either the asymptotic covariance matrix of $\hat{\theta}$ or bootstrap methods.[12]

For the interpretation of results it is important to note that (1.6) optimizes a utility function and not a measure of the distance between forecasted and realized returns. Therefore, θ can be found relevant for one characteristic even if it conveys no information at all about expected returns. The characteristic may just be a predictor of a currency's contribution to the overall skewness or kurtosis of the portfolio, for example. Conversely, a characteristic may be found irrelevant for investment purposes even if it does help in forecasting returns. Indeed, it may forecast

[12] We use bootstrap methods for standard errors in the empirical part of this paper, as these are slightly more conservative and do not rely on asymptotic results.

both higher returns and higher risk for a currency, offering a trade-off that is irrelevant for the investor's utility function.

Menkhoff, Sarno, Schmeling, and Schrimpf (2011b) show that momentum strategies incur higher transaction costs than the carry trade. They even find that momentum profits are of little relevance in currencies of developed countries after transaction costs. So one valid concern is whether the gains of combining momentum with carry persist after taking into consideration time and cross-currency variation in transaction costs. Fortunately, parametric portfolio policies can easily incorporate transaction costs that vary across currencies and over time. This is a particularly appealing feature of the method, since transaction costs varied substantially as foreign exchange trading shifted towards electronic crossing networks.

To address this issue we optimize:

$$\hat{\theta} = \arg\max_{\theta} \frac{1}{T} \sum_{t=0}^{T-1} U \left(rf_t^{US} + \sum_{i=1}^{N_t} (\theta^T x_{i,t}/N_t) r_{t+1}^i - \sum_{i=1}^{N_t} \left| \theta^T x_{i,t}/N_t \right| c_{i,t} \right) \tag{1.7}$$

where $c_{i,t}$ is the transaction cost of currency i at time t, which we calculate as:

$$c_{i,t} = \frac{F_{t,t+1}^{ask} - F_{t,t+1}^{bid}}{F_{t,t+1}^{ask} + F_{t,t+1}^{bid}} \tag{1.8}$$

This is one half of the bid-ask spread as a percentage of the mid-quote.

This assumes the investor buys (sells) a currency in the forward market at the ask (bid) price, and the forward is settled at the next month's spot rate. This may overstate transaction costs. For instance, Mancini, Ranaldo, and Wrampelmeyer (2011) document that effective costs in the spot market are less than half those implied by bid-ask quotes as there is significant within-quote trading.

There is another important point to highlight about transaction costs: for a given month and currency, these are proportional to the absolute weight put on that particular currency. This absolute weight is a function of all the currency characteristics as seen in equation 1.2, so transaction costs will depend crucially on the time-varying interaction between characteristics. One example is the interaction between momentum and other characteristics. As Grundy and Martin (2001) show for stocks, the way momentum portfolios are built guarantees time-varying interaction with other stock characteristics. For instance, after a bear market, winners tend to be low-beta stocks and the reverse for losers. So the momentum portfolio, long in previous winners and short in previous losers, will have a negative beta. The opposite holds after a bull market. The same applies for currencies, after a period where carry experienced high returns, high yielding currencies tend to have positive momentum. In this case, momentum reinforces the carry signal and results in larger absolute weights and thus higher transaction costs. However, after negative carry returns the opposite happens: high yielding currencies have negative momentum. So momentum partially offsets the carry signal resulting in smaller absolute weights and actually reduces the overall transaction costs of the portfolio. This means the transaction costs of including momentum for an extended period of time in a diversified portfolio policy will be lower than what one finds examining momentum in isolation as in Menkhoff, Sarno, Schmeling, and Schrimpf (2011b).

1.3 Empirical analysis

As figure 1.1 shows, combining reversal and momentum with the carry trade considerably mitigated the crash of the carry trade in the last quarter of 2008. Yet this is easy to point out *ex post*. The relevant question is whether investors in the currency market had reasons to believe in the virtue of diversifying their investment strategy before the 2008 crash. For example, Levich and Pojarliev (2011) examine a sample of currency managers and find that they explored carry, momentum and value strategies before the crisis but shifted substantially across investment styles over time. In particular, right before the height of the financial crisis in the last quarter of 2008, most currency managers were heavily exposed to the carry trade, neutral on momentum and investing *against* value. This raises the question of whether the benefits of diversification were as clear before the crisis as they later became apparent. Equally weighting carry, momentum, and value was not an obvious strategy at the time. This also shows that what appear to be naively simple strategies such as equal weighting carry, momentum, and value are not naive at all and in fact benefit a lot from hindsight.

To address this issue we conduct two tests: i) a pre-sample test with the first 20 years of data up to 1996 to determine which characteristics were relevant back then; ii) an out-of-sample experiment since 1996 in which the investor chooses the weight to put on each signal using only historical information available up to each moment in time.

Section 1.3.1. explains the data sources and the variables used in our optimization. In section 1.3.2. we conduct the pre-sample test with the sample from 1976:02 to 1996:02. In section 1.3.3. we conduct the out-of-sample experiment of portfolio optimization using only the relevant variables identified in the pre-sample test.

1.3.1 Data

We use data on exchange rates, the forward discount / premium, and the real exchange rate for the Euro zone and 27 member countries of the Organization for Cooperation and Development (OECD). The countries in the sample are: Australia, Austria, Belgium, Canada, the Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Japan, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Slovakia, South Korea, Spain, Sweden, Switzerland, the UK, and the US.

The exchange rate data are from Datastream. They include spot exchange rates at monthly frequency from November 1960 to December 2011 and one-month forward exchange rates from February 1976. As in Burnside, Eichenbaum, Kleshchelski, and Rebelo (2011) we merge two datasets of forward exchange rates (against the USD and the GBP) to have a comprehensive sample of returns in the forward market in the floating exchange rate era.[13]

We calculate the real exchange rates of each currency against the USD using the spot exchange rates and the consumer price index. The Consumer Price Index (CPI) data come from the OECD/Main Economic Indicators (MEI) online database. For the Euro, the series that starts January 1996 was extended back to January 1988 using the weights of the Euro founding members. In the case of Australia, New Zealand, and Ireland (before November 1975) only quarterly data are available. In those cases, the value of the last available period was carried forward to

[13] The first dataset has data on forward exchange rates (bid and ask quotes) against the GBP from 1976 to 1996 and the second dataset has the same information for quotes against the USD from 1996 to 2011.

the next month.

We test the economic relevance of carry, momentum, and value proxies combined in a currency market investment strategy. The variables used in the optimization exercise are:

1. $\mathbf{sign}_{i,t}$: The sign of the forward discount of a currency with respect to the USD. It is 1 if the foreign currency is trading at a discount ($F_{i,t} > S_{i,t}$) and -1 if it trades at a premium. This is the carry trade strategy examined in Burnside, Eichenbaum, and Rebelo (2008), Burnside (2011), Burnside, Eichenbaum, Kleshchelski, and Rebelo (2011). Given the extensive study of this strategy we adopt it as the benchmark throughout the analysis.

2. $\mathbf{fd}_{i,t}$: The interest rate spread or the forward discount on the currency. We standardize the forward discount using the cross-section mean and standard deviation across all countries available at time t, μ_{FD_t} and σ_{FD_t} respectively. Specifically, denoting the (unstandardized) forward discount as $FD_{i,t}$, we obtain the standardized discount as: $fd_{i,t} = \frac{FD_{i,t} - \mu_{FD_t}}{\sigma_{FD_t}}$. This cross-sectional standardization measures the forward discount in standard deviations above or below the average across all countries. By construction, a variable standardized in the cross-section will have zero mean, implying that the strategy is neutral in terms of the base currency (the US

dollar). Jurek (2009) shows that an interest rate spread strategy similar to this outperforms the equally-weighted carry trade based on sign.

3. $\mathbf{mom}_{i,t}$: For currency momentum we use the cumulative currency appreciation in the last three-month period, cross-sectionally standardized. This variable explores the short-term persistence in currency returns. We use momentum in the previous three months because there is ample evidence for persistence in returns for portfolios with this formation period while there are no significant gains (in fact the momentum effect is often smaller) considering longer formation periods (see Menkhoff, Sarno, Schmeling, and Schrimpf (2011b)). Three-month momentum was also used in Kroencke, Schindler, and Schrimpf (2011). Cross-sectional standardizations means that momentum measures relative performance. Even if all currencies fall relative to the USD those that fall less will have positive momentum.

4. $\mathbf{rev}_{i,t}$: Long-term reversal is the cumulative real currency depreciation in the previous five years, standardized cross-sectionally. First we calculate the cumulative real depreciation of currency i between the basis period (h) and moment t as an index num-

ber $Q_{i,h,t} = \frac{S_{i,t} CPI_{i,h-2} CPI^{US}_{t-2}}{S_{i,h} CPI_{i,t-2} CPI^{US}_{h-2}}$. We use a two-month lag to ensure the CPI is known. We pick $h = t - 60$ which corresponds to 5 years. Then we standardize $Q_{i,h,t}$ cross-sectionally to obtain $rev_{i,t}$. This is essentially the same as the notion of "currency value" used in Asness, Moskowitz, and Pederson (2009). We just use the cumulative deviation from purchasing power parity, instead of the cumulative return as they did, to obtain a longer out-of-sample test period. Reversal is positive for those currencies that experienced the larger real depreciations against the USD in the previous 5 years and negative for the others.

5. $\mathbf{q}_{i,t}$: The real exchange rate standardized by its historical mean and standard deviation. First, as for reversal, we compute $Q_{i,h_i,t}$ with the difference that here the basis period (h_i) is the first month for which there is CPI and exchange rate data available for currency i. Then we compute $q_{i,t} = \frac{Q_{i,h_i,t} - \overline{Q}_{i,t}}{\sigma_{Q_{i,t}}}$, where $\overline{Q}_{i,t}$ is the historical average $\sum_{j=h_i}^{t} Q_{i,h_i,j}/t$ and $\sigma_{Q_{i,t}}$ is the historical standard deviation $\sigma\left(\{Q_{i,h_i,j}\}_{j=h_i}^{t}\right)$. The real exchange rate is measured in standard deviations above or below the historical average. Jordà and Taylor (2009) also used the demeaned real exchange rate but our time series standardization ensures only information available up to each moment in time is used. Unlike rev, which is cross-sectionally

standardized, q is not neutral in terms of the basis currency (the USD). It will tend to be positive for all currencies when these are undervalued against the USD by historical standards.

6. $\mathbf{ca}_{i,t}$: The current account of the foreign economy as a percentage of Gross Domestic Product (GDP), standardized cross-sectionally. The optimization assumes that the previous year current account information becomes known in April of the current year. The current account data were retrieved from the Annual Macroeconomic database of the European Commission (AMECO), where data are available on a yearly frequency from 1960 onward. Many studies examine the relation between the current account and exchange rates justifying its inclusion as a conditional variable.[14]

In order to be considered for the trading strategies, a currency must satisfy three criteria: i) there must be ten previous years of real exchange rate data; ii) current forward and spot exchange quotes must be available; and iii) the country must be an OECD member in the period considered. After filtering out missing observations, there are a minimum of 13 and a maximum of 21 currencies in the sample. On average there are 16 currencies in the sample.

[14] See, for example, Dornbusch and Fischer (1980), Obstfeld and Rogoff (2005), Gourinchas and Rey (2007).

1.3.2 Pre-sample results

Table 1.1 shows the investment performance of the optimized strategies from 1976:02 to 1996:02. We use this pre-sample period to check which variables had strong enough evidence supporting their relevance back in 1996, before starting the out-of-sample experiment.

The two versions of the carry trade ($sign$ and fd) deliver similar performance, with high Sharpe ratios (0.96 and 0.99, respectively) but also with significant crash risk (as captured by excess kurtosis and left-skewness). Momentum provides a Sharpe ratio of 0.56, better than the performance of the stock market of 0.40 in the same sample. This confirms the results of Okunev and White (2003), Burnside, Eichenbaum and Rebelo (2011), and Menkhoff, Sarno, Schmeling, and Schrimpf (2011b).

Financial predictors work better in our optimization than fundamentals like the real exchange rate and the current account. Reversal had an interesting Sharpe ratio of 0.36. This confirms the results of Menkhoff, Sarno, Schmeling, and Schrimpf (2011b) and Asness, Moskowitz, and Pederson (2009). The strategies using the current account and the real exchange rate as conditioning variables achieved modest Sharpe ratios (of 0.16 and 0.07), not at all impressive – especially as this is an in-sample optimization.[15]

The seventh row shows the performance of an optimal strategy combining the carry (both $sign$ and fd) with momentum and reversal – all the statistically relevant variables. Already in 1996 there was ample evidence indicating that a strategy combining different variables lead to substantial gains. The Sharpe ratio of the optimal strategy was nearly 40% higher than the benchmark and it produced a 16.43 percentage points gain in annual certainty equivalent.

Adding fundamentals to this strategy does not improve it: the Sharpe ratio increases only 0.01 and the annual certainty equivalent only 13 basis points. An insignificant gain since in-sample any additional variable must always increase utility.

Table 1.2 shows the statistical significance of the variables, isolated and in combination. The table presents the point-estimates of the coefficients and the bootstrapped p-values (in brackets). We perform the bootstrap by generating 1,000 random samples drawn with replacement from the original sample and with the same number of observations (240 months of returns and respective conditional variables). Then we find the optimal coefficients in each random sample, thereby obtaining their

[15] We also tested these variables out-of-sample (although, based on the in-sample evidence, the investor would choose not to consider them) and found that they did not add to the economic value of the strategy.

distribution across samples.

Taken in isolation, the carry trade variables (*sign* and fd) and momentum are all significant at the 1% level. Reversal has a p-value of 5.3%.

The current account and the real exchange rate have the wrong sign (underweighting undervalued currencies and those with strong current accounts) but are not significant. We have known since Meese and Rogoff (1983), currency spot rates are nearly unpredictable by fundamentals. Using time-series methods, Gourinchas and Rey (2007) find that the current account forecasts the spot exchange rate of the US dollar against a basket of currencies.[16] But we find no evidence in the cross section that the current account is relevant for designing a profitable portfolio of currencies. At best, the fundamental information is subsumed by interest rates, momentum and reversal.

Combining all variables confirms our main result. Carry, momentum and reversal are relevant for the optimization, fundamentals are not. The final row shows the results for an optimization using only the variables deemed relevant. The p-values show the four variables contribute significantly to the economic value of the strategy in combination.

Concerning both carry variables (*sign* and fd), the correlation of

[16] Gourinchas and Rey (2007) derive their result making a different use of the current account information. Namely, they detrend it and also consider net foreign wealth.

their returns was 0.46 from 1976:02 to 1996:02, a value that has not changed much since. So these two ways of implementing the carry trade are not identical and the investor finds it optimal to use both. The *sign* variable assigns the same weight to a currency yielding 0.1% more than the USD as to another yielding 5% more. In contrast, the fd variable assigns weights proportionally to the magnitude of the interest rate differential. Whenever the USD interest rate is close to the extremes of cross section, the *sign* is very exposed to variations in its value, while fd is always dollar-neutral.

One word of caution on forward-looking bias is needed here. Our in-sample test shows that in 1996 some of the strategies recently proposed in the literature on currency returns would already be found to have an interesting performance. This is a necessary condition to assess if investors would want to use these variables in real time to build diversified currency portfolios. However, this does not tell us whether there were other investment approaches that would have seemed relevant in 1996 and resulted afterwards in poor economic performance.

1.3.3 Out-of-sample results

We perform an out-of-sample (OOS) experiment to test the robustness of the optimal portfolio combining carry, momentum, and value strategies. The first optimal parametric portfolio is estimated using the initial 240 months of the sample. Then the model is re-estimated every month, using an expanding window of data, until the end of the sample. The out-of-sample returns thus obtained minimize the problem of look-ahead bias. We do not use q and ca in the optimization as these failed to pass the in-sample test with data until 1996.[17]

The in-sample results also hold out of sample. Table 1.3 shows that the model using interest rate variables, momentum and reversal achieves a certainty equivalent gain of 10.84 percent over the benchmark, with better kurtosis and skewness. Its Sharpe ratio is 1.15, a gain of 0.45 over the benchmark *sign* portfolio.

Transaction costs can considerably hamper the performance of an investment strategy. For example, Jegadeesh and Titman (1993) provide compelling evidence that there is momentum in stock prices, but Lesmond et al. (2004) find that after taking transaction costs into consideration there are little to no gains to be obtained in exploiting momentum.

Panel B of table 1.3 shows the OOS performance of the strategies after taking transaction costs into consideration. Clearly transaction costs matter. The Sharpe ratio of the optimal strategy is reduced by 0.29, a magnitude similar to the equity premium, and the certainty

[17] Although including these does not change much the results as they receive little weight in the optimization.

equivalent drops from 18.87 percent to just 12.15 percent. Momentum and reversal individually show no profitability at all after transaction costs. This finding mirrors the results of Lesmond et al. (2004) with regard to stock momentum. It also confirms the result in Menkhoff, Sarno, Schmeling, and Schrimpf (2011b) that there are no significant momentum profits in currencies of developed countries after transaction costs.

But we find that transaction costs can be managed. In panel C we adjust the optimization to currency and time-specific transaction costs. We calculate the cost-adjusted interest rate spread variable as: $\widetilde{FD}_{i,t} = sign(FD_{it})(|FD_{it}| - c_{it})$ and standardize it in the cross-section to get $\widetilde{fd}_{it}$. We then model the parametric weight function as:

$$w_{i,t} = I(c_{it} < |FD_{it}|)\left[\theta^T x_{i,t}/N_t\right] \tag{1.9}$$

where $I(.)$ is the indicator function, with a value of one if the condition holds and zero otherwise. We maximize expected utility with this new portfolio policy, estimating θ after consideration of transaction costs.

This method effectively eliminates from the sample currencies with prohibitive transaction costs and reduces the exposure to those that have a high ratio of cost to forward discount. Other, more complex, rules might lead to better results, but we refrain from this pursuit as this simple approach is enough to prove the point that managing transaction costs adds considerable value.

The procedure increases the Sharpe ratio of the diversified strategy from 0.86 to 1.06 and produces a gain in the certainty equivalent of 4.54 percent per year. This gain alone is higher than the momentum or reversal certainty equivalents per se. Indeed, the performance of the diversified strategy with managed transaction costs is very close to the strategy in panel A without transaction costs.

Managing transaction costs is particularly important as these currency strategies are leveraged. Given the high Sharpe ratios attainable by investing in currencies, the optimization picks high levels of leverage. We define leverage as $L_t = \sum_{i=1}^{N_t} |w_{it}|$. This is the absolute value of US dollars risked in the currency strategy per dollar invested in the risk-free asset. The optimal strategy has a mean leverage of 5.94 in the OOS period of 1996:03 to 2011:12. As a result, a small difference in transaction costs can have a large impact in the economic performance of the strategy.

One concern in optimized portfolios is whether in-sample overfitting leads to unstable and erratic coefficients OOS. Figure 1.2 shows the estimated coefficients of the diversified portfolio with managed costs in the OOS period. The coefficients of the four variables used are very stable, leading to consistent exposure to the conditioning variables.

The optimal diversified portfolio has a robust OOS economic performance. In the next section we compare it with simple strategies proposed in the literature.

1.4 Comparison with naive currency strategies

We want to assess the importance of using our optimization procedure by comparing our strategy with simple alternatives. This is especially important because DeMiguel, Garlappi, and Uppal (2009) show that simple rules of investment have robust out-of-sample performance when compared to optimized portfolios. One could argue though that simple currency strategies are not so naive. The performance of long-short portfolios depends on the characteristic used to sort currencies in the first place. The choice of characteristics to average is thus crucial. Why carry, momentum and reversal and not something else? There is the choice of designing a strategy that is neutral in terms of the basis currency (as fd) or not (as $sign$). The weighing of different currency characteristics is also arbitrary in a naive strategy. So the scope for arbitrary choices influenced by ex post observation of the data is not necessarily small for naive strategies. Still, the simple strategies found in the literature provide a natural benchmark for our optimal portfolio policy.

We compare the economic performance of the optimal diversified strategy with 5 simple portfolios: i) the *sign* strategy, which is long currencies yielding more than the USD and short the others; ii) the version of momentum (mom_b) proposed in Burnside, Eichenbaum, and Rebelo (2011) which is long currencies with a positive return in the previous month and short the others; iii) an equal-weighted combination of sign and momentum; iv) the interest rate spread strategy (fd); v) an equal-weighted portfolio of the signals used in our portfolio policy – momentum, reversal, *sign* and fd.

It is questionable whether the EW strategy is a naive approach since this strategy uses the signals selected by the optimized portfolio. But including this EW portfolio allows an assessment of how relevant it is to manage transaction costs and to allow the coefficients in the strategy to differ from equality.

Table 1.4 shows the economic performance of the optimal strategy compared to the simple alternatives. All strategies include a 100% investment in the risk free asset complemented with a long-short currency portfolio. We scale all simple strategies to have constant leverage throughout the period, set to match the mean leverage of the optimized strategy. This ensures that differences in performance do not depend on differences in leverage. Note that the leverage of the portfolio is optimally chosen indirectly in the maximization of the utility function. The leverage ($L_t = \sum_{i=1}^{N_t} |w_{it}| = \sum_{i=1}^{N_t} |\theta x_{it}|/N_t$) depends both on the estimated coefficients and on the level of the explanatory variables and therefore changes through time. We also include a version of the optimal strategy with constant leverage to assess if time-varying leverage is important to performance.

The optimal strategy, with a Sharpe ratio of 1.06 and a certainty equivalent of 16.69 percent, outperforms all others. The 0.22 gain in Sharpe ratio with respect to the EW strategy (the 'naive' approach that performed the best) is statistically significant with a p-value of 0.027.[18]This is because the optimal coefficients are not equal (as seen in Figure 1.2) and the simple strategy does not manage transaction costs. The gain in certainty equivalent of 7.13 percentage points is even more expressive.

Perhaps surprising is the unimpressive performance of the combination of *sign* and mom_b. It achieves a lower Sharpe ratio than the *sign* strategy alone. This is because leverage is set to a constant level, so the outperformance of this strategy documented in Burnside, Eichenbaum, and Rebelo (2011) comes from time-varying leverage. Whenever a currency yields more than the USD but experiences a negative return

[18]Computed with same method as DeMiguel, Garlappi, and Uppal (2009).

in the previous month, the two signals cancel out resulting in a weight of zero for the currency. As a result, the combination of $sign$ and mom_b has time-varying leverage, increasing after months when carry has positive returns and decreasing otherwise.

The optimal strategy with constant leverage has a good performance, with a Sharpe ratio of 0.99, though not as good as the unconstrained strategy. Allowing leverage to change over time leads to lower kurtosis and less negative skewness.

All in all, the evidence on economic performance is clear: the optimal strategy produces a certainty equivalent gain of 7.13 percentage points per year over the best performing naive strategy. This gain is due to a higher Sharpe ratio and lower crash risk (as captured by kurtosis and left-skewness).

In table 1.5 we regress the excess returns of the optimal strategy on those of the simple portfolios to assess its abnormal returns, captured by the intercept. The t-statistics and R-squares are obviously significant, since the optimal strategy is built with similar variables as the naive strategies. But these variables do not fully explain the excess returns which range from 0.68 to 2.28 percent per month. The optimal strategy shows an abnormal return of 8.16 percent per year with respect to the best performing naive strategy.

Figure 1.3 shows the cumulative excess returns of each naive strategy compared to the optimal diversified portfolio. We also include the excess return on the stock market portfolio for comparison. Currency strategies in general outperform the stock market. The Sharpe ratio of the stock market in the OOS period is 0.29, lower than any currency strategy examined.

But the graph also shows that no simple portfolio systematically outperforms the optimal strategy. This contrasts with the result of DeMiguel, Garlappi, and Uppal (2009) for stocks. This result extends and confirms recent findings that optimization methods can outperform more naive approaches in currency markets (Corte, Sarno, Tsiakas (2009), Berge, Jordà, and Taylor (2010)).

1.5 Risk exposures

Cochrane (2011) uses the expression "factor zoo" to describe the growing number of risk factors proposed in the literature to explain asset returns. The literature on currency markets is no exception and many sets of risk factors have been proposed, mostly to explain the returns of the carry trade.

Lustig, Roussanov, and Verdelhan (2011a) propose an empirically-motivated high-minus-low factor of currencies sorted on interest rates

(HML_{FX}) to explain carry trade returns. This is an approach similar in spirit to the Fama and French (1992) three-factor model for stock returns. Note however that the HML_{FX} factor is itself by construction a carry portfolio. So while this approach establishes that there is systematic risk in the carry trade, it does not provide intuition on what is the fundamental risk source that justifies its returns. Brunnermeier, Nagel, and Pederson (2008) argue that liquidity-risk spirals are the source of risk of the carry trade. They use the innovation in the TED spread and in the VIX as factors proxying for liquidity and risk. Menkhoff, Sarno, Schmeling, and Schrimpf (2011a) propose innovations in foreign exchange market volatility as a risk factor to explain the carry trade and currency momentum. They also use the innovation in average transaction costs and argue the information in this is subsumed by FX volatility. Lustig and Verdelhan (2007) and Lustig, Roussanov, and Verdelhan (2011b) propose consumption growth risk as a factor to explain the carry returns. Table 1.6 shows the exposure of the optimal diversified strategy (with managed transaction costs) to 8 sets of risk factors.

The first model shows that the currency strategy is not exposed to consumption growth risk.[19] This confirms the results of Burnside (2011)

[19] For this we use the monthly growth rate of Real Personal Consumption Expenditures downloaded from the Federal Reserve of St. Louis.

and Jordà and Taylor (2011).

The second and third models show that our strategy is exposed to liquidity risk (as captured by innovations in the TED spread) and increases in stock volatility (as captured by the changes in VIX). The VIX is a more significant variable, its beta has a t-statistic of -3.98 versus -2.90 for the TED spread.

The fourth model regresses the returns of the optimal strategy on innovations in transaction costs (the cross-section average in the forward exchange market). This does not yield significant results as the adjusted R-squared is negative.

The fifth model shows the diversified portfolio, with a t-statistic of -2.15, is exposed to innovations in foreign exchange volatility confirming Menkhoff, Sarno, Schmeling, and Schrimpf (2011a).[20] But the adjusted R-squared is only 1.88, much less than the 7.27 of the VIX.

Our optimal strategy is also somewhat exposed to stock market risk as the CAPM and the Carhart (1997) models show. But the only relevant variable is the excess return on the market portfolio with a t-statistic of 4.02 in the CAPM and 4.08 in the Carhart four-factor model.

The best performing model, in term of adjusted R-squared, is the

[20] We follow Menkhoff, Sarno, Schmeling, and Schrimpf (2011a) in computing FX volatility in month t as: $\sigma_{FX,t} = \frac{1}{D_t} \sum_{\tau=1}^{D_t} \sum_{i=1}^{N_\tau} \frac{|r_{i,\tau}|}{N_\tau}$, where D_t is the number of trading days in month t and N_τ is the number of currencies available in day τ.

empirically-motivated HML_{FX} factor of Lustig, Roussanov, and Verdelhan (2011a). In this model we regress the optimal portfolio excess returns on RX (the dollar-return of an equal-weighted average of all currency portfolios) and HML_{FX}, the difference in return between the highest yielding currencies and the lowest yielding currencies.[21] The beta with respect to the HML_{FX} is clearly significant, with a t-stat of 6.54, and the adjusted R-squared of 20.85 is by far the highest among the eight models used.

But the most striking result is the consistently high α of the optimal strategy, ranging between 1.73 and 2.38 percent per month, always significant at conventional levels of confidence. So, while the optimal strategy is exposed to some of the factors proposed in the literature on currency returns, the R-squared is typically low and the abnormal returns highly significant.

There is evidence of time-varying risk exposures in the carry trade (Christiansen, Ranaldo, and Söderllind (2010)). In particular, the exposure of the carry to the stock market rises after shocks to liquidity and risk. This is not captured by the unconditional analysis in table 1.6. So it is of interest to ask whether the optimal strategy also has time-varying risk.

[21] We retrieve the data from Adrien Verdelhan's webpage. The data is for returns with all currencies and after transaction costs.

Following Christiansen, Ranaldo, and Söderllind (2010) we run the following OLS regression:

$$r_{p,t} - rf_t = \alpha + \beta_0 RMRF_t + \beta_1 RMRF_t z_{t-1} + \beta_2 R_{bonds,t} + \beta_3 R_{bonds,t} z_{t-1} + \varepsilon_t \tag{1.10}$$

where z_{t-1} is a proxy for (lagged) risk and $R_{bonds,t}$ is the excess return of the 10 year US bond over the risk-free rate.[22] As proxies for risk we use the foreign exchange volatility, the TED spread, VIX, the average transaction cost, and leverage. The first four are also used in Christiansen, Ranaldo, and Söderllind (2010). We add leverage as this is time varying in the optimal strategy and could naturally induce time-varying risk.

The results of the regression are in table 1.7. The only interaction term that is significant is for the TED spread with the market. But the sign of the coefficient is negative, implying the strategy is less exposed to the stock market after a liquidity squeeze. In order for time-varying risk to explain the returns of the diversified strategy, the opposite should happen. All other interaction terms are not significant, so time-varying risk is of little relevance to explain the performance of the diversified strategy. In particular, there is no evidence that the optimal strategy is riskier when it is more leveraged. In general, the conditional models do not add much to the CAPM, and the large significant α persists after

[22] Bond returns are from Datastream.

considering time-varying risk.

Either unconditionally or conditionally the risk factors proposed to explain the carry trade can do very little to explain the returns of our optimal diversified currency strategy. This indicates that the optimal strategy exploits market inefficiencies rather than loading on factor risk premiums.

1.6 Value to diversified investors

We assess whether the currency strategies are relevant for investors already exposed to the major asset classes. Indeed, there is no reason *a priori* that investors should restrict themselves to pure currency strategies, particularly when there are other risk factors that have consistently offered significant premiums as well.

The value of currency strategies to diversified investors holding bonds and stocks is a relatively unexplored topic. Most of the literature on the currency market has focused on currency-specific strategies. One exception is Kroencke, Schindler, and Schrimpf (2011) who find that combining investments in stocks and bonds with currencies improves the Sharpe ratio from 0.34 to 0.43 without entailing an increase in crash risk.

We continue to assume that the investor optimizes power utility with constant relative risk aversion of 4. The returns on wealth are now:

$$R_{p,t+1} = rf_t^{US} + \sum_{j=1}^{M} w_j F_j + \sum_{i=1}^{N_t} w_{i,t} r_{i,t+1} - \sum_{i=1}^{N_t} |w_{i,t}| \, c_{i,t} \qquad (1.11)$$

where w_j are the (constant) weights on a set of M investable factors F expressed as excess returns, and $w_{i,t}$ depends on the characteristics and the θ coefficients that maximize utility jointly with w_j.

Table 1.8 shows the OOS performance of the portfolios with and without the currency strategy. The currency strategy combines the interest rate spread, sign, momentum, and long-term reversal. Subsequently, each two rows compare a portfolio of investable factors with a portfolio combining these factors with the currency strategy.

The opportunity to invest in currencies is clearly valuable to investors. Including currencies in the portfolio always adds to the Sharpe ratio and raises the certainty equivalent. The OOS gains in certainty equivalent range between 9.99 percentage points for an investment in stocks and bonds and 38.04 percentage points for a diversified investment using the Carhart factors. The gain with respect to the Carhart factors comes mainly from the dismal performance of stock momentum in 2009, when it experienced one of its worst crashes in history (Daniel and Moskowitz (2011)).

These gains are far more impressive than the gains from adding factors like HML and SMB to the stock market. Indeed, only the inclusion of bonds improves upon the certainty equivalent of the stock market OOS. Generally, the inclusion of SMB, HML, and WML factors improves Sharpe ratios, but this increase is offset by higher drawdowns, resulting in lower certainty equivalents.

Including currencies however leads to substantial gains. This extends the evidence in Burnside (2011) that there is no known set of risk factors that prices currency and stock returns simultaneously. The relevance of the interest rate spread, currency momentum, and long-term reversal to forecast currency returns makes all conventional risk premiums seem small in comparison.

Including currencies in the portfolio of stocks and bonds produces increases in the Sharpe ratio as high as 0.81 for a portfolio of US stocks and currencies. On average adding currency strategies increases the Sharpe ratio by 0.51. This confirms the results of Kroencke, Schindler, and Schrimpf (2011).

One possible justification for the higher Sharpe ratios obtainable by investing in currencies is that these might entail a higher crash risk – as Brunnermeier, Nagel, and Pedersen (2008) shows for the carry trade. But diversified currency strategies do not conform to this explanation.

Figure 1.4 shows how complementing a portfolio policy with investments in the currency market contributes to performance, including kurtosis and skewness. The currency strategies increase Sharpe ratios and certainty equivalents and, most notably, they also reduce substantially the excess kurtosis and left-skewness of diversified portfolios.

Our results make it hard to reconcile the economic value of currency investing with the existence of some set of risk factors that drives returns in currencies and other asset classes. The substantial increases in Sharpe ratios combined with the lower crash risk indicate that there is either a specific set of risk factors in the currency market or that currency returns have been anomalous throughout our sample.

1.7 Speculative capital

We cannot justify the profitability of our currency strategy as compensation for risk. The obvious alternative explanation is market inefficiency. This might arise due to insufficient arbitrage capital, possibly because strategies exploring the cross section of currency returns were not well known. Jylhä and Suominen (2011) find carry returns explain hedge fund returns controlling for the other factors proposed by Fung and Hsieh (2004) and that growth in hedge fund speculative capital is driving carry trade profits down.

Following Jylhä and Suominen (2011), we run an OLS regression of the returns of the optimal strategy on hedge fund assets under management scaled by the monetary aggregate M2 of the 11 currencies in their sample ($AUM/M2$) and new fund flows ($\Delta AUM/M2$).[23] The regression uses the out-of-sample returns, after transaction costs, of the optimal strategy from 1996:03 to 2008:12 as the dependent variable. The estimated coefficients (and t-statistics in parenthesis) are:

$$\begin{array}{ccccccc} r_{p,t} & = & 0.08 & -1.47 & (\frac{AUM}{M2})_{t-1} & +3.56 & (\Delta\frac{AUM}{M2})_t \\ & & (4.29) & (-3.23) & & (0.36) & \end{array}$$

The new flow of capital to hedge funds is not significant in the regression but the estimated coefficient has the correct sign. The level of hedge fund capital predicts negatively the returns of the optimal strategy. With a t-statistic of -3.23, this provides convincing evidence that the returns of the diversified currency strategy are an anomaly that is gradually being corrected as more hedge fund capital exploits it.

This opens the question whether the large returns of the strategy are likely to continue going forward. We note that in the last three years of our sample (2009-2011) the strategy produces a Sharpe ratio of 0.82, lower than its historical average but still an impressive performance (though not much different than the stock market in the same period).

[23] We thank Matti Suominen for providing us the time series of AUM/M2. See their paper for a more detailed description of the data.

1.8 Conclusion

Diversified currency investments using the information of momentum, yield differential, and reversal, outperform the carry trade substantially. This outperformance materializes in a higher Sharpe ratio and in less severe drawdowns, as reversal and momentum had large positive returns when the carry trade crashed. The performance of our optimal currency strategy poses a problem to peso explanations of currency returns.

Our optimal currency portfolio picks stable coefficients for the relevant currency characteristics and, by dealing with transaction costs, outperforms naive benchmarks proposed in the literature.

The economic performance of the optimal currency portfolio cannot be explained by risk factors or time-varying risk. This suggests market inefficiency or, at least, that the right risk factors to explain currency momentum and reversal returns have not been identified yet. Investing in currencies significantly improves the performance of diversified portfolios already exposed to stocks and bonds. So currencies either offer exposure to some set of unknown risk factors or have anomalous returns.

The most convincing explanation for the returns of our optimal diversified currency portfolio is that it constitutes an anomaly – one which is being gradually arbitraged away as speculative capital increases in the foreign exchange market.

Strategy	Max	Min	Mean	std	kurt	skew	SR	CE
fd	15.91	-25.39	19.23	19.47	4.48	-1.23	0.99	18.97
mom	17.31	-11.60	8.01	14.23	1.80	0.21	0.56	11.61
rev	8.38	-11.31	3.09	8.72	2.24	-0.26	0.36	8.95
sign	21.29	-30.11	17.96	18.74	7.35	-0.90	0.96	18.29
ca	2.79	-3.47	0.61	3.86	1.24	-0.44	0.16	7.59
q	2.02	-2.32	0.12	1.79	4.44	-0.74	0.07	7.34
fd, mom, rev, sign	56.83	-32.78	44.30	32.89	5.54	0.66	1.35	34.72
All	60.38	-25.56	45.28	33.70	5.10	0.60	1.34	34.85

Table 1.1: The in-sample performance of the investment strategies in the period 1976:02 to 1996:02. The optimizations use a power utility with CRRA of 4. The mean, standard deviation and Sharpe ratio are annualized and "Kurt." stands for excess kurtosis

fd	mom	rev	sign	ca	q
7,93	-	-	-	-	-
[0, 000]	-	-	-	-	-
-	5,10	-	-	-	-
-	[0, 009]	-	-	-	-
-	-	3,46	-	-	-
-	-	[0, 053]	-	-	-
-	-	-	3,71	-	-
-	-	-	[0, 000]	-	-
-	-	-	-	-2,19	-
-	-	-	-	[0, 247]	-
-	-	-	-	-	-0,15
-	-	-	-	-	[0, 648]
4,75	8,10	6,57	2,67	-4,95	-1,08
[0, 051]	[0, 019]	[0, 078]	[0, 085]	[0, 621]	[0, 828]
7,09	7,26	4,37	2,44	-	-
[0, 004]	[0, 006]	[0, 051]	[0, 023]	-	-

Table 1.2: The statistical significance of the variables in the in-sample period of 1976:02 to 1996:02. The coefficient estimates and bootstrapped p-values (in brackets).

	Max	Min	Mean	std	kurt	skew	SR	CE
Panel A: No transaction costs								
fd	18,64	-29,20	21,38	24,33	2,15	-0,82	0,88	10,89
mom	14,72	-10,03	4,97	13,29	0,57	0,04	0,37	4,39
rev	9,42	-9,67	1,69	9,50	1,42	0,23	0,18	2,84
sign	16,40	-21,21	15,01	21,37	1,95	-0,64	0,70	8,03
all in	26,90	-22,88	38,02	32,98	0,12	-0,14	1,15	18,87
Panel B: With transaction costs								
fd	4,59	-10,92	2,80	7,40	5,01	-1,35	0,38	4,55
mom	0,64	-1,33	-0,02	0,66	17,41	-2,43	-0,03	2,88
rev	0,00	0,00	0,00	0,00	18,62	2,06	0,05	2,91
sign	12,12	-16,30	8,89	15,70	2,14	-0,67	0,57	6,59
all in	20,39	-18,31	19,20	22,20	0,54	-0,16	0,86	12,15
Panel C: With $w_{i,t} = I(c_{it} < \|FD_{it}\|) \left[\theta^T x_{i,t}/N_t\right]$								
fd	12,83	-20,70	11,91	17,18	2,66	-0,89	0,69	8,35
mom	6,67	-7,01	2,14	6,04	2,37	-0,07	0,35	4,33
rev	3,44	-3,84	-0,37	3,00	4,66	-0,16	-0,12	2,36
sign	18,10	-23,09	12,08	20,23	2,74	-0,76	0,60	5,98
all in	26,70	-22,75	28,48	26,84	0,69	-0,16	1,06	16,69

Table 1.3: The OOS performance of the investment strategies in the period 1996:03 to 2011:12 with different methods to deal with transaction costs. Panel A presents the results without considering transaction costs. Panel B takes transaction costs into consideration. Panel C excludes all currencies whenever the bid-ask spread is higher than the forward discount, then adjusts the forward discount by the transaction cost. All optimizations use a power utility function with a CRRA of 4 and the coefficients are re-estimated each month using an expanding window of observations in the OOS period of 1996:03 to 2011:12.

Strategy	Max	Min	Mean	std	kurt	skew	SR	CE
sign	24.72	-34.04	19.23	32.09	2.06	-0.61	0.60	-2.00
mom_b	63.26	-40.43	14.60	40.69	4.45	0.53	0.36	-17.53
$sign$+mom_b	55.78	-51.58	21.32	45.34	3.31	0.02	0.47	-27.72
fd	20.57	-29.16	17.54	23.77	2.73	-0.64	0.74	7.81
EW$(sign, fd, mom, rev)$	19.69	-25.77	22.76	27.09	1.17	-0.60	0.84	9.56
constant leverage	30.06	-22.39	27.73	27.89	0.92	-0.22	0.99	14.51
Optimal strategy	26.70	-22.75	28.48	26.84	0.69	-0.16	1.06	16.69

Table 1.4: The performance of naive portfolios in the OOS period compared to the optimal strategy using sign, fd, momentum and reversal. The naive strategies have a constant leverage of 5.94, the same as the optimal strategy on average. The OOS returns are from 1996:03 to 2011:12. The optimization uses an expanding window of returns, re-estimating the coefficients each month. Results with transaction costs.

Strategy	α	t-stat	β	t-stat	R^2
sign	1.39	3.58	0.62	14.92	54.20
mom_b	2.28	4.05	0.08	1.57	1.30
sign+mom_b	1.89	3.74	0.27	7.10	21.17
fd	1.43	3.02	0.65	9.56	32.71
EW$(sign, fd, mom, rev)$	0.68	2.72	0.89	28.85	81.57

Table 1.5: The OOS performance of the optimal strategy regressed on the naive portfolios. The regressions are standard OLS regressions. The optimal strategy uses sign, fd, momentum and reversal and re-estimates the coefficients in the OOS period every month. Results after transaction costs. The alphas are expressed in percentage points per month. The OOS returns are from 1996:03 to 2011:12.

α	β_1	β_2	β_3	β_4	Adj-Rsquared
$r_{opt,t} = \alpha + \beta_1 \Delta cons_t + \varepsilon_t$					
2.08	1.31	-	-	-	-0.10
[3.20]	[0.90]	-	-	-	-
$r_{opt,t} = \alpha + \beta_1 \Delta TED_t + \varepsilon_t$					
2.38	-0.06	-	-	-	3.76
[4.31]	$[-2.90]$	-	-	-	-
$r_{opt,t} = \alpha + \beta_1 \Delta VIX_t + \varepsilon_t$					
2.38	-0.46	-	-	-	7.27
[4.41]	$[-3.98]$	-	-	-	-
$r_{opt,t} = \alpha + \beta_1 \Delta c_t + \varepsilon_t$					
2.37	26.80	-	-	-	-0.42
[4.21]	[0.45]	-	-	-	-
$r_{opt,t} = \alpha + \beta_1 \Delta \sigma_{FX.t} + \varepsilon_t$					
2.38	-8.71	-	-	-	1.88
[4.28]	$[-2.15]$	-	-	-	-
$r_{opt,t} = \alpha + \beta_1 RMRF_t + \varepsilon_t$					
2.19	0.44	-	-	-	7.43
[4.03]	[4.02]	-	-	-	-
$r_{opt,t} = \alpha + \beta_1 RMRF_t + \beta_2 SMB_2 + \beta_3 HML_t + \beta_4 WML_t + \varepsilon_t$					
2.07	0.50	0.01	0.20	0.09	6.88
[3.74]	[4.08]	[0.05]	[1.19]	[0.87]	-
$r_{opt.t} = \alpha + \beta_1 RX_t + \beta_2 HML_{FX.t} + \varepsilon_t$					
1.73	0.32	1.35	-	-	20.85
[3.37]	[1.16]	[6.54]	-	-	-

Table 1.6: Risk exposures of the optimal strategy. We regress the OOS returns of the optimal strategy (after transaction costs) on each set of risk factors. Standard OLS coefficients (and t-statistics in brackets). The OOS returns are from 1996:03 to 2011:12.

z	α	$RMRF_t$	$RMRF_t z_{t-1}$	$R_{bonds,t}$	$R_{bonds,t} z_{t-1}$	Adj-rsquared
σ_{FX}	2.26	0.46	-0.07	-0.22	0.00	6.70
	[3.95]	[3.73]	[-0.85]	[-0.77]	[0.01]	-
TED	2.09	0.52	-0.15	-0.04	-0.19	9.31
	[3.69]	[4.33]	[-2.35]	[-0.14]	[-1.39]	-
VIX	2.26	0.52	-0.15	-0.15	-0.07	7.51
	[3.99]	[3.99]	[-1.60]	[-0.53]	[-0.51]	-
c	2.22	0.46	-0.08	-0.18	0.01	6.77
	[3.85]	[3.79]	[-1.04]	[-0.63]	[0.06]	-
leverage	2.20	0.46	0.15	-0.18	0.04	7.07
	[3.84]	[3.88]	[1.29]	[-0.71]	[0.14]	-

Table 1.7: Time-varying risk of the optimal strategy. In each row we regress the OOS returns, after transaction costs, of the optimal strategy on the market and bond returns, using a different risk proxy as a state variable to account for time-varying risk exposure. We standardize all risk proxies subtracting the mean and dividing by the standard deviation. Standard OLS coefficients and t-statistics (in brackets). The optimal strategy uses sign, fd, momentum and reversal and re-estimates the coefficients in the OOS period every month. The OOS returns are from 1996:03 to 2011:12.

Strategy	Max	Min	Mean	STD	Kurt.	Skew	SR	CE
fd, mom, rev and sign	26,70	-22,75	28,48	26,84	0,69	-0,16	1,06	16,69
Stock market	7,16	-14,94	3,17	12,46	1,36	-0,81	0,25	2,83
Stock market+curr.	27,27	-21,87	27,95	26,93	0,73	-0,16	1,04	16,07
FF factors	19,89	-29,96	12,94	27,41	1,53	-0,84	0,47	-1,53
FF factors+curr.	31,75	-22,26	27,06	25,79	1,32	0,13	1,05	16,83
Carhart factors	33,51	-63,23	20,84	35,67	8,02	-1,37	0,58	-30,46
Carhart factors+curr.	19,29	-24,21	15,78	22,92	0,94	-0,45	0,69	7,58
Stocks and bonds	8,13	-13,74	5,39	12,16	2,15	-0,93	0,44	5,19
Stock and bonds+curr.	23,53	-22,67	27,98	27,47	0,80	-0,28	1,02	15,19

Table 1.8: The OOS performance of portfolios combining a currency strategy with different background assets. The currency strategy uses momentum, the interest rate spread, reversal and sign. Each row denoted with '+curr.' combines the available factors with the currency strategy. Results with transaction costs. Optimizations carried out with a CRRA of 4 and 240 months in the initial in-sample estimate. The OOS period is from 1996:03 to 2011:12.

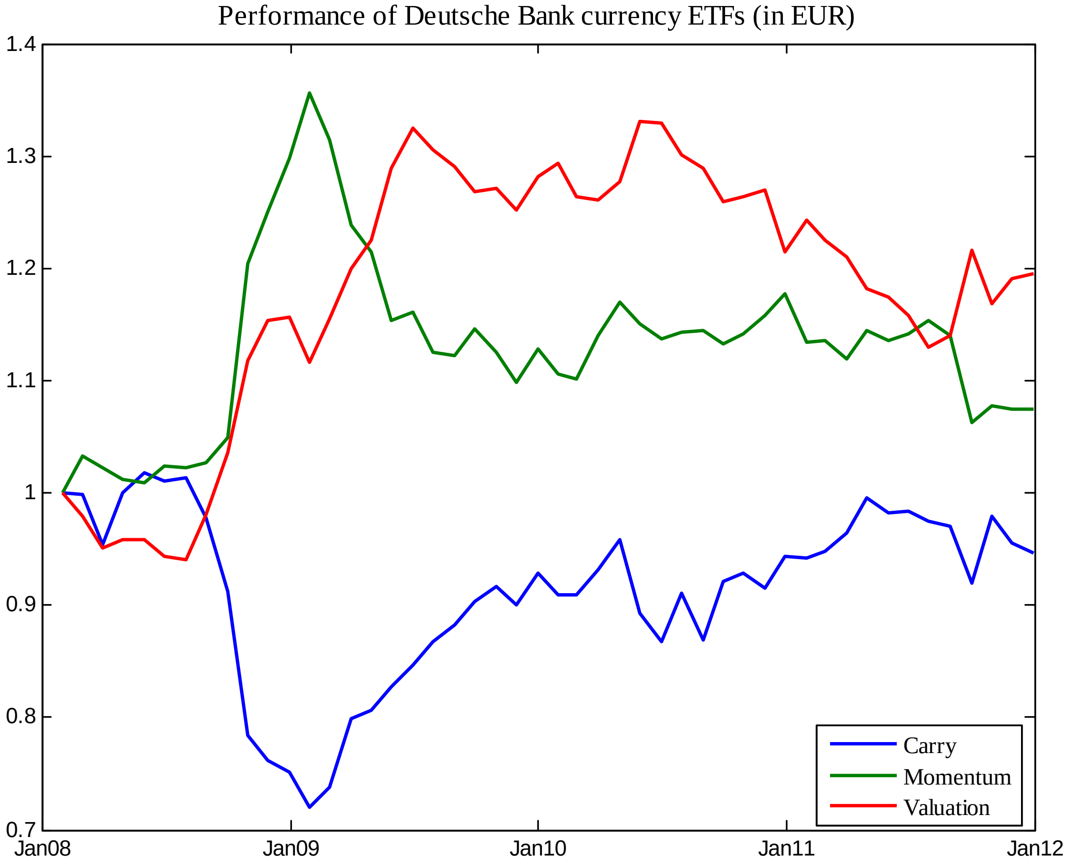

Figure 1.1: The performance of Deutsche Bank currency ETFs (in euros). Each line plots the cumulative monthly returns of a Deutsche Bank ETF from 2008:01 to 2011:12.

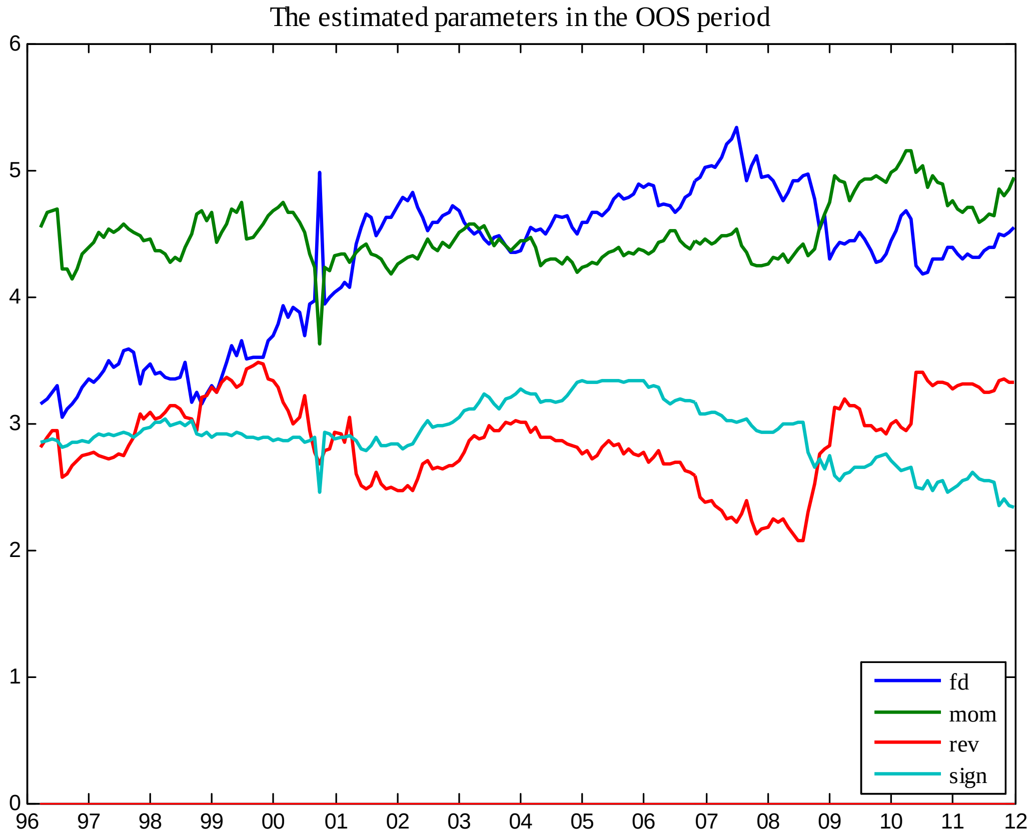

Figure 1.2: The estimates of the coefficients of the portfolio in the OOS period from 1996:03 to 2011:12. Optimization with CRRA of 4 and considering transaction costs.

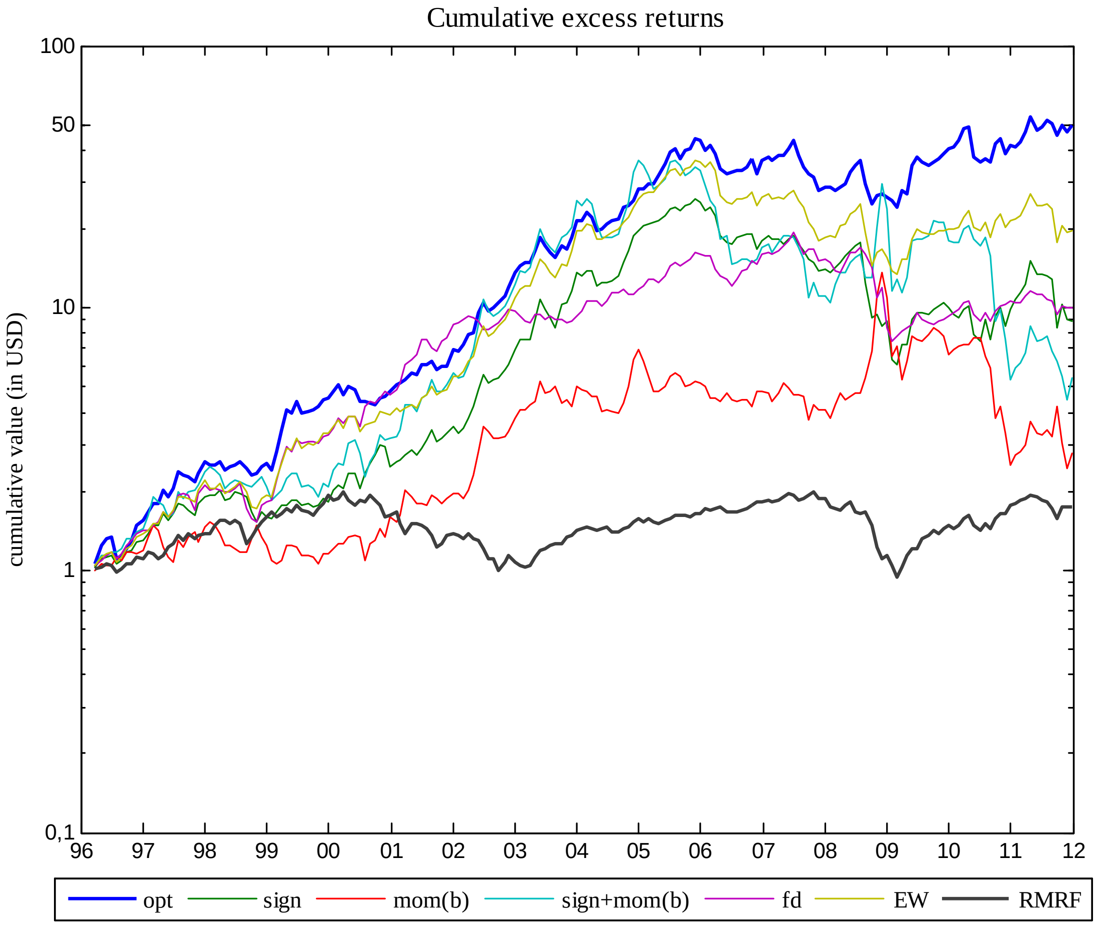

Figure 1.3: OOS performance of the strategies versus naive portfolios in the period of 1996:03 to 2011:12. The naive currency strategies have a constant leverage of 5.94 to match the mean leverage of the optimized strategy in the OOS period. Returns after transaction costs.

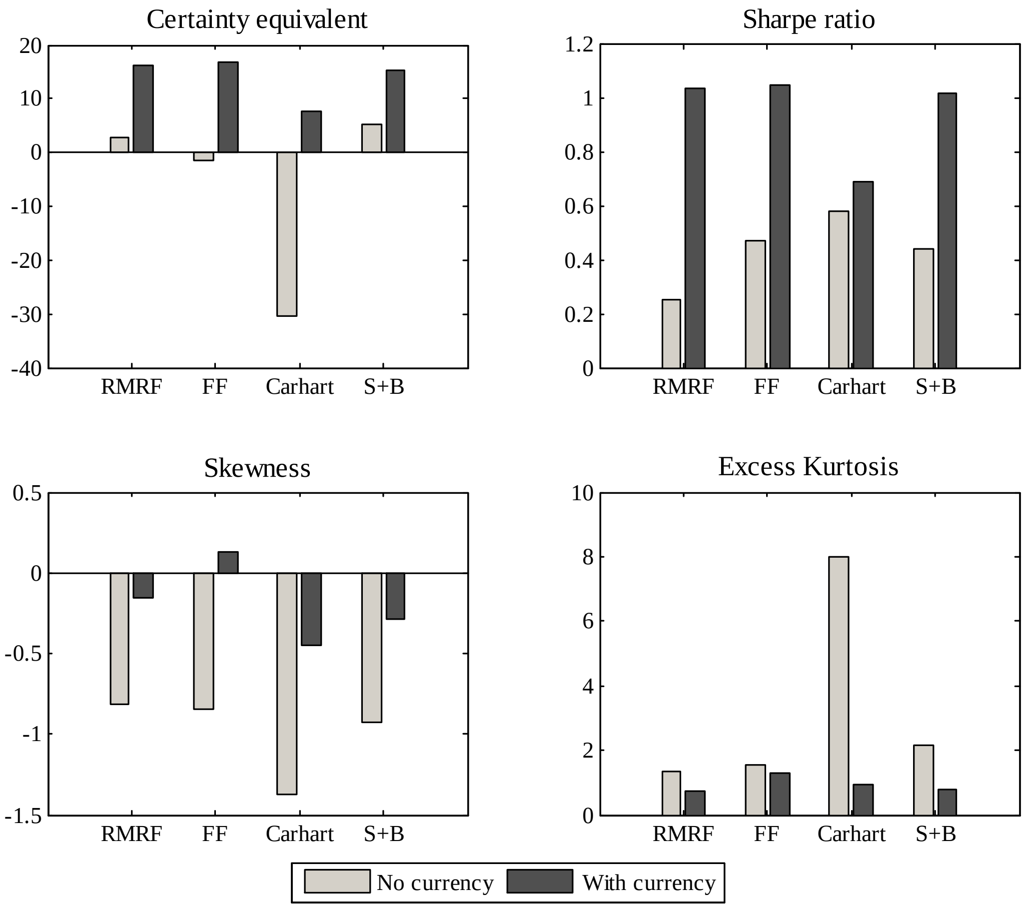

Figure 1.4: The OOS value of currency strategies for investors exposed to different background risks. Each set of columns shows the performance of an optimized portfolio with the available assets (light grey) and one which combines it with the currency strategy (dark grey). The currency strategy uses the information on the interest rate spread, sign, momentum and reversal. The OOS period is from 1996:03 to 2011:12. Results with transaction costs.

Chapter 2

The Bottom-up Beta of Momentum

2.1 Introduction

Unconditionally, the Fama-French factors do not explain the risk or returns of momentum (Fama and French (1996)). But Grundy and Martin (2001) argue this is because momentum portfolios have time-varying systematic risk, which is not captured in unconditional regressions. The winners-minus-losers beta should depend by construction on the previous returns of the market.

For instance, in late 1999, and after good returns in the stock market, the winners tended naturally to be high beta stocks while the laggards should mainly be low beta stocks. Hence the momentum portfolio, short on previous losers and long on previous winners, should have a high beta by design. By contrast, at the end of 2008, in an extreme bear market, previous losers should be typically stocks with high betas such as financials, while the group of winner stocks would have low betas. Thus the momentum portfolio would have a negative beta by construction.

But Grundy and Martin (2001) did not actually examined if the composition of the momentum portfolio changes over time as their conjecture suggests. For that one needs to estimate the betas of individual stocks at each point in time and from these obtain the aggregate beta of the portfolio from the bottom up.

In this work, we compute the bottom-up beta of momentum using monthly returns from January 1950 to December 2012 for all stocks in the Center of Research for Security Prices (CRSP). This allows a direct test of Grundy and Martin's (2001) conjecture. We find the bottom-up beta changes quite substantially over time, ranging from a minimum of -1.71 to a maximum of 2.09 and that it is positively related to previous returns in the market. The conjecture is thus confirmed.

Bottom-up betas are much better at explaining the risk of momentum than an unconditional regression. The bottom-up betas with respect to the Fama-French factors explain 39.59 percent of the out-of-sample variation in momentum returns, nearly 17 times more than one unconditional regression.

The bottom-up betas do not explain the alpha of momentum though, which remains significantly positive. This confirms that it is not persistence in the returns of the Fama-French factors that explain momentum profits (Grundy and Martin (2001), Blitz, Huij, and Martens (2011)).

Using bottom-up betas to hedge the risk of momentum in real time does not avoid its large drawdowns. The hedged strategy still has a high excess kurtosis and a pronounced left-skew. This confirms Daniel and Moskowitz (2011) who also find momentum investors could not use time-varying betas to avoid the crashes in real time.

This chapter is organized as follows. Section 2.2, explains the estimation method of the conditional betas with respect to the CAPM model and discuss whether these can explain the risk and returns of momentum. Section 2.3 extends the analysis to the Fama-French factors. Section 2.4 shows the performance of the hedged momentum strategies and section 2.5 concludes.

2.2 The time-varying beta of momentum

The momentum portfolio changes its composition as new stocks join the group of previous losers or the group of previous winners. This changing composition should induce time variation in the beta of the portfolio with respect to the market. We compare three methods of estimation of these time-varying betas: i) a bottom-up approach; ii) estimating beta as a linear function of factor returns in the formation period; iii) a high-frequency beta estimated from daily returns.

To estimate the bottom-up betas, we use data from CRSP (Center for Research in Security Prices) with monthly returns for all stocks listed in the NYSE, AMEX or NASDAQ from January 1950 to December 2010. Following the standard practice, the momentum portfolios are sorted according to accumulated returns in the formation period which is from month t-12 to month $t-2$. The stocks are classified into deciles using as cutoff points the universe of all firms listed on the NYSE. This way there is an equal number of firms listed in the NYSE in each decile. This is to prevent the possibility of very small firms dominating either the long or short leg of the portfolio.

In order to be considered in the portfolio a firm's stock must have a valid return in month $t-2$, a valid price in month $t-13$, and information on the market capitalization of the firm in the previous month. We take into consideration the delisting return of a stock whenever it is available. Individual stocks are value-weighted within each decile. The return of the winner-minus-losers (WML) is simply the return of the top decile portfolio, sorted on previous momentum, minus the return of the bottom decile portfolio.

We estimate the beta of each individual stock running an OLS regression of its monthly excess return on the excess return of the market from $t-61$ to $t-2$, the end of the formation period. We require at least 24 valid returns in that period to estimate the beta. The market return is the value-weighted return of all stocks in the CRSP universe, as obtained from Kenneth French's online data library.

The bottom-up beta of the momentum portfolio is the weighted average of individual betas in the portfolio:

$$\beta_{BU,t} = \sum_{i=1}^{N_t} w_{i,t} \hat{\beta}_{i,t} \tag{2.1}$$

where N_t is the number of stocks in the portfolio at time t, $w_{i,t}$ is the weight of stock i in the WML portfolio and $\hat{\beta}_{i,t}$ is the beta of the individual stock estimated from past monthly returns. This beta relies only on past information, known before time t, so its forecasts are out-of-sample (OOS) by construction.

Figure 2.1 shows the bottom-up beta of the momentum strategy and compares it with the unconditional beta, obtained from running a regression of the WML on the market with the full sample.

The unconditional beta is -0.27. So on average the losers portfolio has a higher beta than the winners portfolio. But this unconditional beta masks substantial time-variation in the composition of the WML portfolio. The bottom-up beta ranges from -1.71 to 2.09. Therefore, the momentum strategy is at times highly exposed to the overall stock market, while at other times it is actually negatively related with the market.

The conjecture of Grundy and Martin (2001) is that the time-varying systematic risk of momentum is due to the return of the overall market during the formation period. After bear markets, winners tend to be low-beta stocks while losers are high-beta stocks. By shorting losers to go long winners, the WML portfolio will have by construction a negative beta. The opposite happens after bull markets. Figure 2.2 shows their conjecture holds true.

Our bottom-up beta, obtained from the individual stock level, is approximately a linear function of market returns during the formation period. Following Grundy and Martin (2001), we estimate this beta running an OLS regression:

$$r_{wml,t} = \alpha + \beta_0 rmrf_t + \beta_1 rmrf_t rmrf_{t-2,t-12} \tag{2.2}$$

where $rmrf_{t-2,t-12}$ is the cumulative gross return of the market during the formation period of the momentum portfolio. Then the time-varying linear beta is:

$$\beta_{L,t} = \beta_0 + \beta_1 rmrf_{t-2,t-12} \tag{2.3}$$

The third method to estimate the time-varying beta of momentum is to use the daily returns of the WML portfolio and those of the market. Following Daniel and Moskowitz (2011) we regress at the end of each month the daily returns of momentum on the market in the previous 126 sessions ($\approx$ 6 months).[1] As the bottom-up beta, the one-month lagged high-frequency beta ($\beta_{HF,t-1}$) produces an OOS forecast of momentum's exposure to the market.

Table 2.1 presents descriptive statistics for the four estimates of beta. The unconditional beta is the estimate from on OLS regression, which is constant. The first two rows show the in-sample results for β_{unc} and β_L using returns from 1964:02 to 2010:12.[2] The third and fourth rows present the OOS results for the same variables. Here we use the period from 1950:01 to 1964:01 to obtain initial estimates of the betas, producing an OOS forecast for the following month. Then we reiterate the procedure every month till the end of the sample using an expanding window of monthly observations. The resulting OOS period is from 1964:02 to 2010:12.

The in-sample (out-of-sample) linear beta varies between a minimum

[1] They estimate the beta using 10 lags of daily returns to correct for stale quotes. This correction does not improve results in my sample period, so we only report results from regressions with no lags.

[2] To facilitate the comparison, the same sample period is examined for all methods. The daily returns of the Fama-French factors is only available starting in 1963:07. This restricts the comparable sample period to start in 1964:02.

of -1.74 (-1.60) and a maximum of 1.61 (1.46). The high-frequency beta varies even more from a minimum of -1.94 to a maximum of 2.16. So all estimates show there is substantial time-variation in the market exposure of the momentum strategy.

The most relevant test is whether time-varying betas explain the risk and returns of momentum. For each estimation method we obtain the hedged momentum return as:

$$z_t = r_{wml,t} - \widehat{\beta}_t rmrf_t \tag{2.4}$$

where $\widehat{\beta}_t$ is the conditional beta at time t.

Table 2.2 shows that time-varying risk does not explain the excess returns of the momentum strategy. The mean excess return of the market-hedged momentum strategy ranges from 1.19 percent per month to 1.47 percent per month. All the t-statistics exceed four, so they are highly significant. This confirms the results of Grundy and Martin (2001) and Blitz, Huij, and Martens (2011). Grundy and Martin (2001) show that time-varying risk does not explain the alpha of momentum. Conversely, Blitz, Huij, and Martens (2011) show that momentum profits come from persistence in returns at the individual stock level, rather than in the factors themselves. As a result, hedging market risk has little effect on the alpha of momentum.

Momentum has more beta risk in 'good times' (in bull markets) than in 'bad times' (in bear markets). The opposite pattern should hold for time-varying beta to explain momentum average excess returns.

But taking time-varying betas into consideration enhances substantially the understanding of momentum's risk.

The r-squared ($1 - \frac{var(z_t)}{var(r_{wml,t})}$) improves OOS from just 1.93 percent for the unconditional model to values ranging from 15.78 percent to 24.81 percent using the conditional models.[3] This is 8 to nearly 13 times more than the unconditional model. The bottom-up beta performs particularly well OOS, with the highest r-squared among those considered.

2.3 Exposure to the Fama-French factors

One unconditional OLS regression of monthly momentum returns on the Fama-French factors from 1964:02 to 2010:12 holds (t-statistics in parenthesis):

$$\begin{array}{rllll} r_{WML,t} = & 1.71 & -\ 0.34\ r_{RMRF,t} & -\ 0.04\ r_{SMB,t} & -\ 0.47\ r_{HML,t} \\ & (5.70) & (-4.91) & (-0.42) & (-4.51) \end{array}$$

[3]Note that the OOS r-squared can assume negative values. Goyal and Welch (2008) show this is often the case with predictive regressions.

The regression has an adjusted r-squared of just 5.63 percent and a significant positive alpha of 1.71 percent, confirming the result in Fama and French (1996) that their factors do not explain the risk and returns of momentum. Still, this improves substantially the fit of the CAPM, which has an adjusted r-squared of just 2.54 percent for the same period. This is mainly because momentum is significantly and negatively related to value. Yet, as for the exposure to the market, these estimates mask substantial time-variation in risk.

Figure 2.3 shows the exposure of momentum to the market (RMRF), value (HML), and size (SMB) factors. Just as for the market, exposure to value and size varies a lot.

Table 2.3 shows the descriptive statistics of the bottom-up betas. The exposure of the momentum portfolio to size and value varies even more than the exposure to the market. For the HML factor, the beta ranges from -3.06 to 2.61, while for the market it ranges only from -1.62 to 1.58. The standard deviation of the betas with respect to size and value are, respectively, 0.82 and 0.83. This is over the double of the 0.40 standard deviation of β_{rmrf}.

Table 2.4 shows the average excess returns of the hedged portfolio with respect to the Fama-French factors, its t-statistics and r-squared.

As for the CAPM, the conditional models do not explain the mean excess returns of momentum which have always a positive mean with significant t-statistics ranging between 3.57 (for the bottom-up beta) and 6.48 (for the beta linear with past returns).

However, the conditional models improve considerably the understanding of the systematic risk of momentum. In sample, the r-squared of the unconditional model is 6.13 percent, which is reduced to only 2.37 percent in the out-of-sample (OOS) test. The high-frequency beta, used in Daniel and Moskowitz (2011), produces an OOS r-squared of 29 percent. In spite of this large improvement, the high frequency approach underperforms other measures of systematic risk. The linear model has an OOS r-squared of 37.72 percent. Even more so, bottom-up betas explain 39.59 percent of the systematic risk of momentum OOS. This is nearly 17 times more than the unconditional model. As for the CAPM, the bottom-up beta is the best conditional model to explain the systematic risk of momentum.

2.4 Systematic risk and momentum crashes

Grundy and Martin (2001) find that hedging the time-varying risk exposures of momentum produces stable returns. Daniel and Moskowitz (2011) show that this relies on using ex post information. Hedging in real time with time-varying betas does not avoid the momentum crashes. However, their method of estimating the time varying risk is with top-down regressions of daily data – the high-frequency beta. We find this is the less satisfactory approach to capture the time-varying systematic risk of momentum (although it still clearly outperforms an unconditional model).

Bottom-up betas provide a superior method to estimate the time-varying exposure of momentum to other factors. This leads to the question of whether hedging this time-varying risk with a more suitable method could avoid the large drawdowns of momentum.

Table 2.5 shows the performance of hedged portfolios using bottom-up betas. Hedging market risk or the Fama-French factor exposures reduces the excess kurtosis and left-skewness of returns, without a clear effect on the Sharpe ratio (it improves using the CAPM but decreases using the Fama-French factors). The reduction in crash risk is modest though. The hedged strategies have an excess kurtosis exceeding 5 and a left-skew almost as pronounced as the WML strategy. This confirms the result of Daniel and Moskowitz (2011) that using time-varying betas does not avoid the crashes in real time. It also confirms the result in chapter 3 that it is not time-varying systematic risk but rather risk specific to momentum that forecasts crashes.

2.5 Conclusion

When the previous returns of a factor are high, the momentum portfolio rotates from low-beta stocks to high-beta stocks on that factor. This changes the betas of momentum over time.

Conditional betas capture the systematic risk of momentum much better than an unconditional model. Using the Fama-French factors, the out-of-sample r-squared increases from just 2.37 percent for the unconditional model to as much as 39.59 percent using conditional betas.

The bottom-up betas perform particularly well in capturing time variation in systematic risk. They achieve the best results comparing to the linear and the high-frequency beta, both with the CAPM and the Fama-French factors.

Using this method to manage the time-varying risk does not avoid momentum crashes though.

Beta	Mean	Max	Min	STD(B)
β_{unc} (IS)	-0.26	-	-	-
β_L (IS)	-0.12	1.61	-1.74	0.52
β_{unc}	-0.08	0.07	-0.24	0.07
β_L	-0.06	1.46	-1.60	0.38
β_{BU}	-0.05	2.09	-1.71	0.57
β_{HF}	0.15	2.16	-1.94	0.62

Table 2.1: Descriptive statistics of different betas of momentum. The first two rows present results for in-sample betas and the others for out-of-sample betas. All betas are from 1964:02 to 2010:12.

Beta	$\overline{z}_t$	$\overline{z}_t/\sigma_{\overline{z}_t}$	R^2
β_{unc} (IS)	1.47	4.95	2.71
β_L (IS)	1.43	5.36	22.14
β_{unc}	1.40	4.70	1.93
β_L	1.37	5.11	20.39
β_{BU}	1.34	5.15	24.81
β_{HF}	1.19	4.32	15.78

Table 2.2: Performance of time-varying betas in explaining excess returns and risk of the momentum strategy. The first two rows present results for in-sample betas and the others for out-of-sample betas. All hedged returns and betas are from 1964:02 to 2010:12.

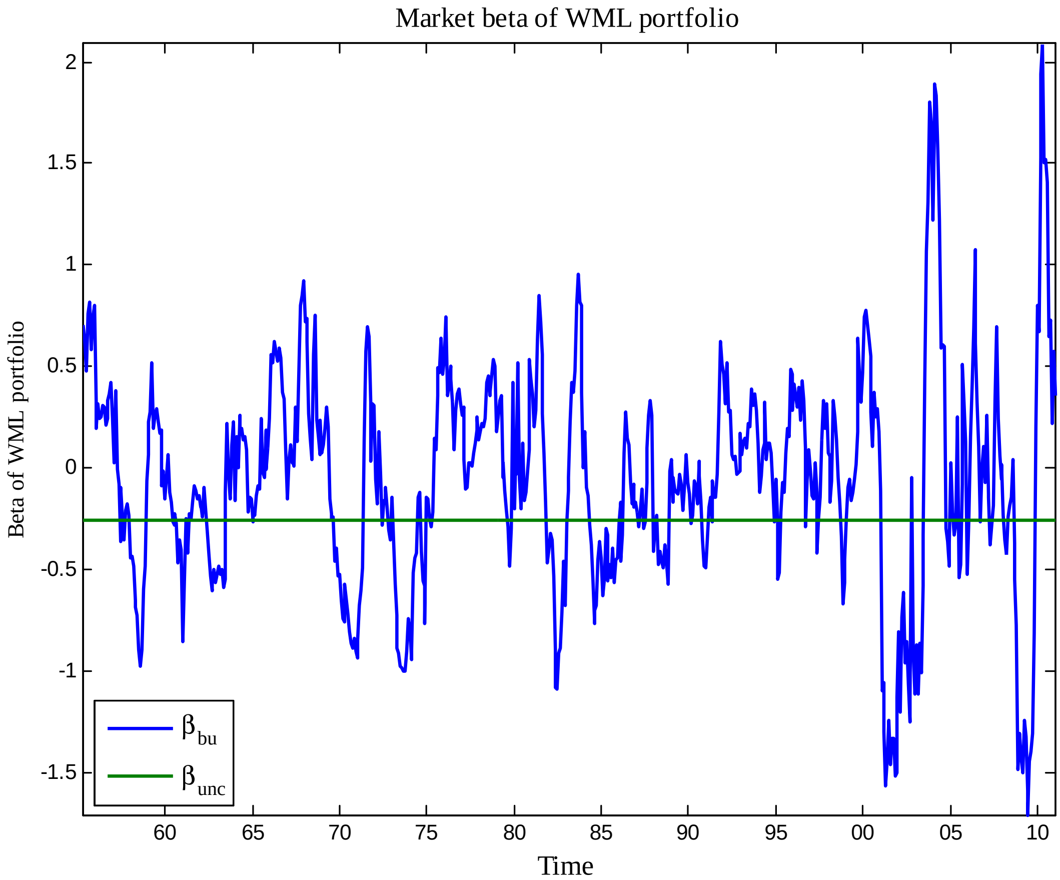

Figure 2.1: Bottom-up and unconditional beta of the WML portfolio. The bottom-up beta of the WML portfolio is obtained from the previous 5 years monthly returns of individual stocks. Returns from 1955:03 to 2010:12.

Beta	Mean	Max	Min	STD(B)
Rm-Rf	-0.03	1.58	-1.62	0.40
SMB	-0.03	2.25	-1.91	0.82
HML	0.07	2.61	-3.06	0.83

Table 2.3: Descriptive statistics of the conditional betas of the WML portfolio (bottom-up) for the Fama-French (1992) model. All betas are from 1955:03 to 2010:12.

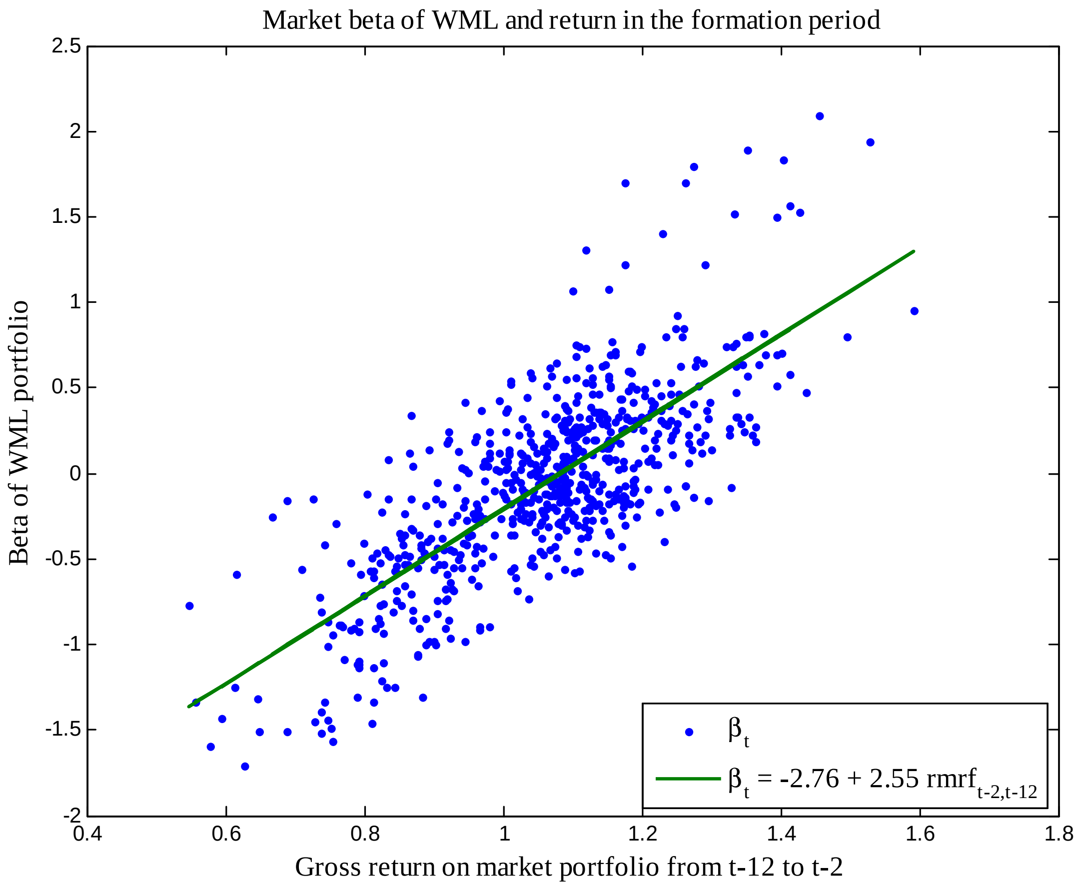

Figure 2.2: The bottom-up market beta of the WML portfolio and the previous return on the market portfolio. All returns from 1955:03 to 2010:12.

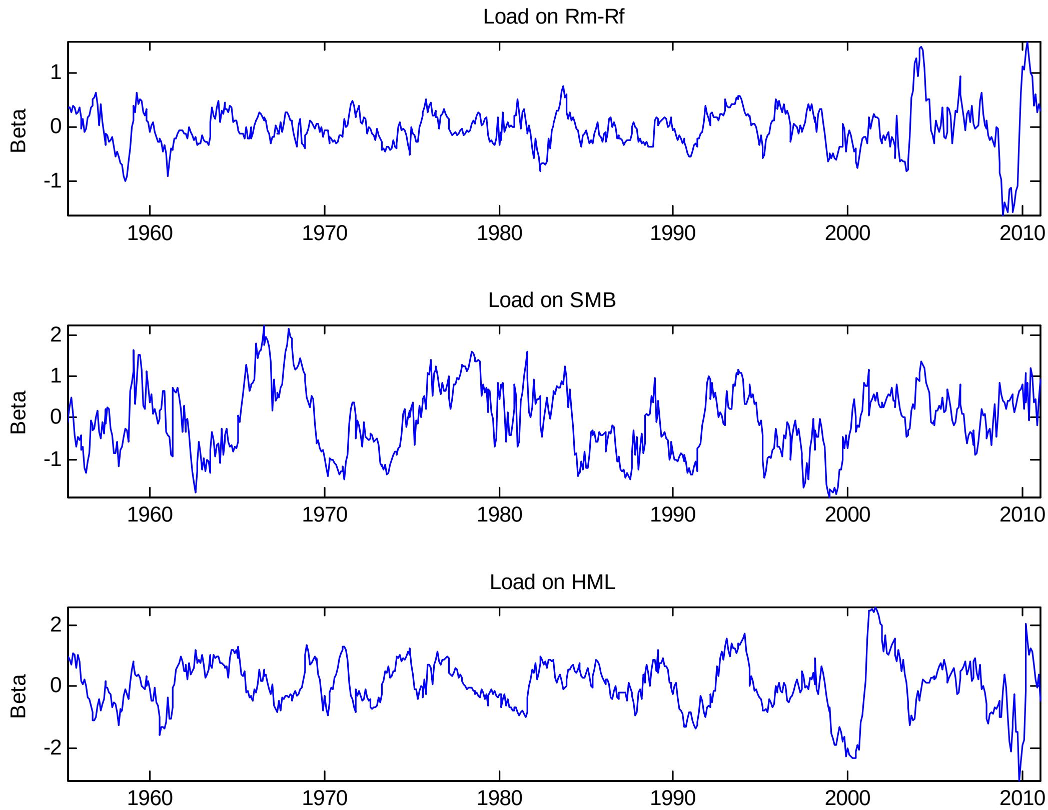

Figure 2.3: The loading of the WML on the FF factors (bottom-up betas). The sample period is from 1955:02 to 2010:02.

Beta	$\overline{z}_t$	$\overline{z}_t/\sigma_{\overline{z}_t}$	R^2
β_{unc} (IS)	1.71	5.85	6.13
β_L (IS)	1.48	6.48	42.73
β_{unc}	1.65	5.56	2.37
β_L	1.39	5.85	37.72
β_{BU}	0.84	3.57	39.59
β_{HF}	1.29	5.1	29.00

Table 2.4: Performance of time-varying betas with respect to the Fama-French factors in explaining excess returns and risk of the momentum strategy. The first two rows present results for in-sample betas and the others for out-of-sample betas. All hedged returns and betas are from 1964:02 to 2010:12.

	Max	Min	Mean	STD	KURT	SKEW	Sharpe
WML	25.54	-45.53	16.14	24.65	7.89	-1.52	0.65
Market hedged	23.41	-38.05	15.98	21.33	5.77	-1.22	0.75
FF hedged	24.07	-37.12	9.87	19.09	6.81	-1.17	0.52

Table 2.5: Performance of the hedged portfolios. In each case the betas are estimated bottom-up. The results are thus OOS. Kurt stands for excess kurtosis. Returns from 1964:02 to 2010:12.

Chapter 3

Managing the Risk of Momentum[0]

3.1 Introduction

Momentum is a pervasive anomaly in asset prices. Jegadeesh and Titman (1993) find that previous winners in the US stock market outperform previous losers by as much as 1.49 percent a month. The Sharpe ratio of this strategy exceeds the Sharpe ratio of the market itself, as well as the size and value anomalies. Momentum returns are even more of a puzzle since they are negatively related to the market and the value risk factors. From 1927 to 2011, momentum had a monthly excess return of 1.75 percent per month controlling for the Fama-French factors. This result has led researchers to use momentum as an additional risk

[0]We thank Eduardo Schwartz for helpful comments and Kent Daniel for letting us use his data.

factor.[1] Momentum is not just a US stock market anomaly. Momentum has been documented in European equities, emerging markets, country stock indices, industry portfolios, currency markets, commodities and across asset classes.[2] Grinblatt and Titman (1989,1993) found most mutual fund managers incorporate momentum of some sort in their investment decisions, so relative strength strategies are widespread among practitioners.

But the remarkable performance of momentum comes with occasional large crashes.[3] In 1932, the winners-minus-losers strategy delivered a -91.59 percent return in just two months. In 2009 momentum experienced another significant crash of -73.42 percent over three months. Even the large returns of momentum do not compensate an investor with reasonable risk version for these sudden crashes that take decades to recover from.

The two most expressive momentum crashes occurred as the market rebounded following large previous declines. One explanation for this pattern is the time-varying systematic risk of the momentum strategy. Grundy and Martin (2001) show that momentum has significant negative

[1] Carhart (1997)

[2] See Rouwenhorst (1998,1999) for international evidence, Asness, Liew, and Stevens (1997) for country stock indices, Moskowitz and Grinblatt (1999) for industry portfolios, Okunev and White (2003) and Menkhof, Sarno, Schmeling, and Schrimpf (2012) for currency markets, Erb and Harvey (2006) for commodities and Asness, Moskowitz, and Pederson (2011) for momentum across asset classes.

[3] Daniel and Moskowitz (2011)

beta following bear markets.[4] They argue that hedging this time-varying market exposure produces stable momentum returns but Daniel and Moskowitz (2011) show that using betas in real time does not avoid the crashes.

In this chapter we propose a different method to manage momentum risk. We estimate the risk of momentum by the realized variance of daily returns and find that it is highly predictable. An auto-regression of monthly realized variances yields an out-of-sample (OOS) R-square of 57.82 percent. This is 19.01 percentage points (p.p.) more than a similar auto-regression for the variance of the market portfolio which is already famously predictable.[5]

Making use of this predictability in risk management leads to substantial economic gains. We scale the long-short portfolio by its realized volatility in the previous 6 months, thereby obtaining a strategy with constant volatility. The Sharpe ratio improves from 0.53 for unmanaged momentum to 0.97 for its risk-managed version. But the most important benefit comes from a reduction in crash risk. The excess kurtosis drops from 18.24 to 2.68 and the left skew improves from -2.47 to -0.42. The minimum one-month return for momentum is -78.96 percent while for

[4] Following negative returns for the overall market, winners tend to be low-beta stocks and the reverse for losers. Therefore the winner-minus-losers strategy will have a negative beta.

[5] Engle and Bollerslev (1986), Schwert (1989).

risk-managed momentum is -28.40 percent. The maximum drawdown of momentum was -96.69 percent versus -45.20 percent for its risk-managed version.

To assess the economic significance of our results, we evaluate the benefits of risk management for a risk-averse investor using a power utility function with Constant Relative Risk Aversion (CRRA) of four. The representative investor holding the market portfolio has an annual certainty equivalent of 0.14 percent. Adding momentum to the portfolio reduces this certainty equivalent to -5.46 percent. However, combining the market with risk-managed momentum already achieves a certainty equivalent of 13.54 percent. We find that the main benefit of risk management comes in the form of smaller crash risk. This alone provides a gain of 14.96 p.p. in annual certainty equivalent.

One pertinent question is why managing risk with realized variances works while using time-varying betas does not. To answer this question we decompose the risk of momentum into systematic risk (from time-varying exposure to the market) and specific risk. We find that the systematic component is only 23 percent of total risk on average, so most of the risk of momentum is specific. This specific risk is more persistent and predictable than the systematic component. The OOS R-square of the specific component is 47 percent versus 21 percent for the systematic component. This is why hedging with time-varying betas fails: it focuses on the smaller part of risk and also the less predictable one.

The work that is more closely related to ours is Grundy and Martin (2001) and Daniel and Moskowitz (2011). But their work focuses on the time-varying systematic risk of momentum, while we focus on momentum's specific risk. Our results have the distinct advantage of offering investors using momentum strategies an effective way to manage risk without forward-looking bias. The resulting risk-managed strategy deepens the puzzle of momentum.

This chapter is structured as follows. Section 3.2 discusses the long-run properties of momentum returns and its exposure to crashes. Section 3.3 shows that momentum risk varies substantially over time in a highly predictable manner. We analyze the implications of such predictability for risk management in Section 3.4. Section 3.5 decomposes the gains of risk management according to the moments of returns. In Section 3.6 we decompose the risk of momentum and study the persistence of each of its components. Finally, Section 3.7 presents our conclusions.

3.2 Momentum in the long run

We compare momentum to the Fama-French factors using a long sample of 85 years of monthly returns from July 1926 to December 2011.[6] This is the same sample period as in Daniel and Moskowitz (2011).

Figure 3.1 presents the cumulative returns of each factor. As each factor consists of a long-short strategy, we construct the series of returns assuming the investor puts $1 in the risk-free asset at the beginning of the sample, buys $1 worth of the long portfolio and sells the same amount of the short portfolio. Then in each subsequent month, the strategy fully reinvests the accumulated wealth in the risk-free asset, assuming a position of this same size in the long and short leg of the portfolio. The winners-minus-losers (WML) strategy offered an impressive performance for investors. One dollar fully reinvested in the momentum strategy grew to $68,741 by the end of the sample. This compares to the cumulative return of $2,136 from simply holding the market portfolio.

There would be nothing puzzling about momentum's returns if they corresponded to a very high exposure to risk. However, running an OLS regression of the WML on the Fama-French factors gives (t-stats in parenthesis):

[6] See the annex for a description of the data.

$$
\begin{array}{llllll}
r_{WML,t} = & 1.752 & -0.378\, r_{RMRF,t} & -0.249\, r_{SMB,t} & -0.677\, r_{HML,t} \\
& (7.93) & (-8.72) & (-3.58) & (-10.76)
\end{array}
$$

so momentum provided abnormal returns of 1.75 percent per month with *negative* exposure to the Fama and French (1992) risk factors. This amounts to a 21 percent per year abnormal return and the negative loadings on other risk factors imply momentum actually diversified risk in this extended sample.

Table 3.1 compares descriptive statistics for momentum in the long-run with the Fama-French factors. Buying winners and shorting losers has provided large returns of 14.46 percent per year, with a Sharpe ratio higher than the market. The impressive excess returns of momentum, its high Sharpe ratio and negative relation to other risk factors, particularly the value premium, make it look like a free lunch for investors.

But as Daniel and Moskowitz (2011) show there is a dark side to momentum. Its large gains come at the expense of a very high excess kurtosis of 18.24 combined with a pronounced left-skew of -2.47. These two features of the distribution of returns of the momentum strategy imply a very fat left tail – significant crash risk. The apparent free lunch of momentum returns can very rapidly turn into a free fall, wiping out decades of returns.

Figure 3.2 shows the performance of momentum in the two most turbulent decades for the strategy: the 1930's and the 2000's.

In July and August 1932, momentum had a cumulative return of -91.59 percent. From March to May 2009, momentum had another large crash of -73.42 percent. These short periods leave an enduring impact on cumulative returns. For example, someone investing one dollar in the WML strategy in July 1932 would only recover it in April 1963, 31 years after and with considerably less real value. This puts the risk to momentum investing in an adequate long-run perspective.

Both in 1932 and in 2009, the crashes happened as the market rebounded after experiencing large losses.[7] This leads to the question of whether investors could predict the crashes in real time and hedge them away.

Grundy and Martin (2001) show that momentum has a substantial time-varying loading on stock market risk. The strategy ranks stocks according to returns during a formation period, for example the previous 12 months. If the stock market performed well in the formation period, winners tend to be high-beta stocks and losers low-beta stocks. So the momentum strategy, by shorting losers to buy winners, has by construction a significant time-varying beta: positive after bull markets and negative after bear markets. They argue that hedging this time-varying

[7]Daniel and Moskowitz (2011) argue this is due to the option-like payoffs of distressed firms in bear markets.

risk produces stable returns, even in pre-WWII data, when momentum performed poorly. In particular, the hedging strategy would be long in the market portfolio whenever momentum has negative betas, hence mitigating the effects of rebounds following bear markets, which is when momentum experiences the worst returns. But the hedging strategy in Grundy and Martin (2001) uses forward looking betas, estimated with information investors did not have in real time. Daniel and Moskowitz (2011) show that using betas estimated solely on ex-ante information does not avoid the crashes and portfolios hedged in real time perform even worse than unhedged momentum.

3.3 The time-varying risk of momentum

One possible cause for the excess kurtosis of momentum is time-varying risk.[8] The very high excess kurtosis of 18.24 of the momentum strategy (more than twice the market portfolio) leads us to study the dynamics of its risk and compare it with the market (RMRF), value (HML) and size (SMB) risk factors.

For each month, we compute the realized variance RV_t from daily returns in the previous 21 sessions. Let $\{r_d\}_{d=1}^{D}$ be the daily returns and $\{d_t\}_{t=1}^{T}$ the time series of the dates of the last trading sessions of each

[8]See, for example, Engle (1982) and Bollerslev (1987).

month. Then the realized variance of factor i in month t is:[9]

$$RV_{i,t} = \sum_{j=0}^{20} r^2_{d_t - j} \tag{3.1}$$

Figure 3.4 shows the monthly realized volatility of momentum. This varies substantially over time, from a minimum of 3.04 percent (annualized) to a maximum of 127.87 percent.

Table 3.2 shows the results of AR (1) regressions with the realized variances of the WML, RMRF, SMB and HML:

$$RV_{i,t} = \alpha + \rho RV_{i,t-1} + \varepsilon_t \tag{3.2}$$

Panel A presents the results for RMRF and WML, for which we have daily data available from 1927:03 to 2011:12. Panel B adds the results for HML and SMB, for which daily data is available from 1963:07 onwards.

Momentum returns are the most volatile. From 1927:03 to 2011:12, the average realized volatility of momentum was 15.03, more than the 12.81 of the market portfolio. For the 1963:07 onwards sample, the average realized volatility of momentum was 16.40, the highest when compared to RMRF, SMB and HML.

In the full sample period, the standard deviation of monthly real-

[9]Correcting for serial correlation of daily returns does not change the results significantly.

ized volatilities is higher for momentum (12.26) than the market (7.82). Panel B confirms this result for the other factors in the 1963:07 onwards sample. So the risk of momentum is the most variable.

The risk of momentum is also the most persistent. The $AR(1)$ coefficient of momentum in the 1963:07 sample is 0.77, which is 0.19 more than the market for the same sample period and higher than the estimate for SMB and HML.

To check the out-of-sample (OOS) predictability of risk, we use a training sample of 240 months to run an initial AR(1) and then use the estimated coefficients and last observation of realized variance to forecast the realized variance in the following month. Then each month we use an expanding window of observations to produce OOS forecasts and compare these to the accuracy of the historical mean $\overline{RV}_{i,t}$. As a measure of goodness of fit we estimate the OOS R-square as:

$$R^2_{i,OOS} = 1 - \frac{\sum_{t=S}^{T-1} (\widehat{\alpha}_t + \widehat{\rho}_t RV_{i,t} - RV_{i,t+1})^2}{\sum_{t=S}^{T-1} (\overline{RV}_{i,t} - RV_{i,t+1})^2} \tag{3.3}$$

where S is the initial training sample, $\widehat{\alpha}_t, \widehat{\rho}_t$ and $\overline{RV}_{i,t}$ are estimated with information available only up to time t.

The last column of table 3.2 shows the OOS R-squares of each auto-regression. The AR(1) of the realized variance of momentum has an

OOS R-square of 57.82 percent (full sample), which is 50 percent more than the market. For the period from 1963:07 to 2011:12, the OOS predictability of momentum-risk is twice that of the market. Hence more than half of the risk of momentum is predictable, the highest level among risk factors. In the next section we explore the potential of this predictability for risk management.

3.4 Risk-managed momentum

We use estimates of momentum risk to scale the exposure to the strategy in order to have constant risk over time.

For each month we compute a variance forecast $\hat{\sigma}_t^2$ from daily returns in the previous 6 months.[10] Let $\{r_{WML,t}\}_{t=1}^{T}$ be the monthly returns of momentum and $\{r_{WML,d}\}_{d=1}^{D}$, $\{d_t\}_{t=1}^{T}$ be, as above, the daily returns and the time series of the dates of the last trading sessions of each month.

The variance forecast is:

$$\hat{\sigma}_t^2 = 21 \sum_{j=0}^{125} r_{d_{t-1}-j}^2 / 126 \qquad (3.4)$$

As WML is a zero-investment and self-financing strategy we can scale it without constraints. We use the forecasted variance to scale the

[10] We also used one-month and three-month realized variances as well as exponentially-weighted moving average (EWMA) with half-lifes of 1, 3 and 6 months. All worked well with nearly identical results.

returns:

$$r_{WML^*,t} = \frac{\sigma_{\text{target}}}{\hat{\sigma}_t} r_{WML,t} \tag{3.5}$$

where $r_{WML,t}$ is the unscaled or plain momentum, $r_{WML^*,t}$ is the scaled or risk-managed momentum, and σ_{target} is a constant corresponding to the target level of volatility. Scaling corresponds to having a weight in the long and short legs that is different from one and varies over time, but it is still a self-financing strategy, so the choice of the constant is arbitrary. We pick a target corresponding to an annualized volatility of 12 percent.[11]

Figure 3.3 shows the cumulative returns of risk-managed momentum compared to plain momentum. The risk-managed momentum strategy achieves a higher cumulative return with less risk. So there are economic gains to risk-management of momentum. The scaled strategy benefits from the large momentum returns when it performs well and effectively shuts it off in turbulent times, thus mitigating momentum crashes. As a result, one dollar invested in risk-managed momentum grows to $6,140,075 by the end of the sample, nearly 90 times more than the plain momentum strategy.[12] Also, the risk-managed strategy

[11] The annualized standard deviation from monthly returns will be higher than 12% as volatilities at daily frequency are not directly comparable to those at lower frequencies due to small positive autocorrelation of daily returns.

[12] This difference in cumulative returns is fundamentally due to risk management

no longer has variable and persistent risk, so risk management indeed works.[13]

Table 3.3 provides a summary of the economic performance of WML^* in 1927-2011. The risk-managed strategy has a higher average return, with a gain of 2.04 (p.p. per year), with substantially less standard deviation (less 10.58 p.p. per year). As a result, the Sharpe ratio of risk-managed portfolios almost doubles from 0.53 to 0.97. The most important gains of risk-management show up in the improvement in the higher order moments. Managing the risk of momentum drops the excess kurtosis from a very high value of 18.24 to just 2.68 and reduces the left skew from -2.47 to -0.42. This practically eliminates the crash risk of momentum. Figure 3.5 shows the density function of momentum and its risk-managed version. Momentum has a very long left tail which is much reduced in its risk-managed version.

The benefits of risk-management are specially important in turbulent times. Figure 3.6 shows the performance of risk-managed momentum in the decades with the most impressive crashes. The scaled momentum

successfully avoiding the two momentum crashes. But in the post-war period from 1946 to 2007, not including the crashes, the Sharpe ratio of momentum was 0.86, versus 1.15 for risk-managed momentum. So risk management also contributes to performance in not so turbulent times.

[13]The AR(1) coefficient of monthly squared returns is only 0.14 for the scaled momentum versus 0.40 for the original momentum. Besides, the auto-correlation of momentum is significant up to 15 lags while only 1 lag for risk-managed momentum. So persistence in risk is much smaller for the risk-managed strategy.

manages to preserve the investment in the 1930's. This compares very favorably to the pure momentum strategy which loses 90 percent in the same period. In the 2000's simple momentum lost 28 percent of wealth, because of the crash in 2009. Risk-managed momentum ends the decade up 88 percent as it not only avoids the crash but also captures part of the positive returns of 2007-2008.

Figure 3.7 shows the weights of the scaled momentum strategy over time – interpreted as the dollar amount in the long or short leg. These range between the values of 0.13 and 2.00, reaching the most significant lows in the early 1930's, in 2000-02, and in 2008-09. On average, the weight is just 0.90, slightly less than full exposure to momentum. As these weights depend only on ex ante information this strategy could actually be implemented in real time.

3.5 Economic Significance: An Investor Perspective

Momentum offers a trade-off between an appealing Sharpe ratio, obtained from the first two moments of its distribution, and less appealing higher order moments, such as high kurtosis and left skewness. An economic criterion is needed to assess whether this trade-off is interesting. Risk management offers improvements to momentum across the board, higher expected returns, lower standard deviation and crash risk. Still it is pertinent to evaluate the relative economic importance of each of these contributions.

We use an utility-based approach to discuss the appeal of momentum to a representative investor. We adopt the power utility function as it has the advantage of taking into consideration higher order moments instead of focusing merely on the mean and standard deviation of returns. This is particularly suitable as momentum has a distribution far from normal. The utility of returns is:

$$U(r) = \frac{(1+r)^{1-\gamma}}{1-\gamma} \tag{3.6}$$

where γ is the Constant Coefficient of Relative risk aversion (CRRA). Bliss and Panigirtzoglou (2004) estimate γ empirically from risk-aversion implicit in one-month options on the S&P and the FTSE and find a value very close to 4. This is a more plausible value for CRRA than previous estimates featured in the equity premium puzzle literature using utility over consumption. So we adopt this value for CRRA. We obtain the certainty equivalent from the utility of returns:

$$CE(r) = \{(1-\gamma)E[U(r)]\}^{\frac{1}{1-\gamma}} - 1 \tag{3.7}$$

This states the welfare a series of returns offers the investor in terms of an equivalent risk-free annual return, expressed in a convenient unit of percentage points (p.p.) per year.

For an economic measure of the importance of the mean return, variance and higher order moments, we use a Taylor series approximation to expected utility around the mean:

$$E[U(r)] = U(\bar{r}) + \frac{1}{2}U''(\bar{r})E(r-\bar{r})^2 + \phi_3(r) \tag{3.8}$$

where ϕ_3 is the Lagrangian rest corresponding to the utility from moments with order greater than 2. From this we obtain the certainty equivalent due to each moment:

$$CE(\mu_1) = \{(1-\gamma)U(\bar{r})\}^{\frac{1}{1-\gamma}} - 1 \tag{3.9}$$

$$CE(\mu_2) = \left\{(1-\gamma)[U(\bar{r}) + \frac{1}{2}U''(\bar{r})E(r-\bar{r})^2]\right\}^{\frac{1}{1-\gamma}} - CE(\mu_1) - 1 \tag{3.10}$$

$$CE(\mu_{i>2}) = CE(r) - CE(\mu_1) - CE(\mu_2) \tag{3.11}$$

We compute the certainty equivalent from annual overlapping returns, an adequate horizon from an investor perspective. Also, one-year horizons capture better the occasional large drawdowns of momentum documented in Section 3.2.

Table 3.4 shows the decomposition of the certainty equivalent for the representative investor holding the market portfolio. It also assesses whether it is optimal to deviate from the market portfolio to include (risk-managed) momentum.

The first row shows the results for holding only the market portfolio. The mean return had a positive contribution for the certainty equivalent of 11.72 percent per year, but the variance of returns reduces this by 7.39 p.p. Higher order moments diminish the certainty equivalent by a further 4.18 p.p. As a result the certainty equivalent of the market portfolio was only 0.14 percent per year.

Adding momentum to the market portfolio increases returns. As a result the certainty equivalent of the mean return increases from 11.72 percent per year to 28.51 percent. The higher standard deviation partially offsets this gain by reducing the certainty equivalent by 6.51 p.p. Still, looking only at the first two moments of the combined portfolio leads to the conclusion that the investor is better off including momentum.

But the increase in higher order risk – the momentum crashes – reduces the certainty equivalent by 15.89 p.p. per year. As a result, including momentum actually dampens the economic performance of the market portfolio. The certainty equivalent of the market plus momentum is -5.46 percent per year versus 0.14 percent of the market only. The high Sharpe ratio of momentum does not compensate the investor for bearing the increased crash risk. So, in spite of the impressive cumulative returns of momentum in the long-run, crash risk is so high that a reasonable risk-averse investor would rather just hold the market portfolio.

This illustrates with an economic measure how far the distribution of momentum is from normality. Indeed, momentum has a distribution with many small gains and few, but extreme, large losses. Taking this into account the momentum puzzle of Jegadeesh and Titman (1993) is substantially diminished. Our discussion in terms of utility shows that rare observations with large losses are significant enough to change the interest of momentum for a risk-averse investor holding the market.

In contrast, risk-managed momentum produces large economic gains. These come from higher returns when compared to the market (a 19.92 p.p. gain) and less crash risk than the market with plain momentum (a 14.96 p.p. gain). As a result, the annual certainty equivalent of the market with risk-managed momentum is 13.54 percent, which compares very favorably to the 0.14 percent of the market alone and even more so with the -5.46 percent of the market combined with simple momentum.

Comparing the strategies with plain and scaled momentum, risk management produces a 3.13 p.p. gain in returns and a 15.86 p.p. gain from reduction of risk. So the bulk of the gains comes from less risk, especially high-order moments. Essentially, scaling momentum eliminates non-normal risk without sacrificing returns.

3.6 Anatomy of momentum risk

A well documented result in the momentum literature is that momentum has time-varying market risk (Grundy and Martin (2001)). This is an intuitive finding since after bear markets winners tend to be low-beta stocks and the inverse for losers. But Daniel and Moskowitz (2011) show that using betas to hedge risk in real time does not work. This contrasts with our finding that the risk of momentum is highly predictable and managing it offers strong gains. Why is scaling with forecasted variances so different from hedging with market betas? We show it is because time-varying betas are not the main source of predictability in momentum risk.

We use a CAPM regression to decompose the risk of momentum into systematic and specific risk:

$$RV_{wml,t} = \beta_t^2 RV_{rmrf,t} + \sigma_{e,t}^2 \tag{3.12}$$

The realized variances and betas are estimated with 6-months of daily returns. On average, the systematic component $\beta_t^2 RV_{rmrf,t}$ accounts for only 23 percent of the total risk of momentum. Almost 80 percent of the momentum risk is specific. Also, the different components do not have the same predictability. Table 3.5 shows the results of an AR (1) on each component of risk.

Either in-sample or out-of-sample (OOS), β_t^2 is the least predictable component of momentum risk. Its OOS R-square is only 5 percent. The realized variance of the market also has a small OOS R-square of 7 percent. When combined, both elements of systematic risk component already show more predictability (OOS R-square of 21 percent) but still less than the realized variance of momentum (OOS R-square of 44 percent). The most predictable component of momentum variance is the specific risk with an OOS R-square of 47 percent, more than double the predictability of the systematic risk.

Hedging with betas alone, as in Daniel and Moskowitz (2011) fails because most of the risk is left out.

3.7 Conclusion

Unconditional momentum has a distribution that is far from normal, with huge crash risk. We find that taking this crash risk into consideration, momentum is not appealing for a risk-averse investor.

Yet, the risk of momentum is highly predictable. Managing this risk eliminates exposure to crashes and increases the Sharpe ratio of the strategy substantially.

As a result, the momentum puzzle shows up stronger in its risk-managed version and the case for a peso explanation of momentum returns is severely weakened.

3.8 Annex: Data sources

We obtain daily and monthly returns for the market portfolio, the high-minus-low, the small-minus-big, the ten momentum-sorted portfolios and the risk-free (one-month Treasury-bill return) from Kenneth French's data library. The monthly data is from July 1926 to December 2011 and the daily data is from July 1963 to December 2011.

For the period from July 1926 to June 1963, we use daily excess returns on the market portfolio (the value-weighted return of all firms on NYSE, AMEX and Nasdaq) from the Center for Research in Security Prices (CRSP). We also have daily returns for ten value-weighted portfolios sorted on previous momentum from Daniel and Moskowitz (2011). This allows us to work with a long sample of daily returns for the winner-minus-losers (WML) strategy from August 1926 to December 2011. We use these daily returns to calculate the realized variances in the previous 21, 63 and 126 sessions at the end of each month.

For the momentum portfolios, all stocks in the NYSE, AMEX and Nasdaq universe are ranked according to returns from month t-12 to t-2, then classified into deciles according to NYSE cutoffs. So there is an equal number of NYSE firms in each bin. The WML strategy consists on shorting the lowest (loser) decile and a long position in the highest (winner) decile. Individual firms are value weighted in each decile. Following the convention in the literature, the formation period for month t excludes the returns in the preceding month. See Daniel and Moskowitz (2011) for a more detailed description of how they build momentum portfolios. The procedures (and results) are very similar to those of the Fama-French momentum portfolios for the 1963:07-2011:12 sample.

	Max	Min	Mean	STD	KURT	SKEW	SR
RMRF	38.27	-29.04	7.33	18.96	7.35	0.17	0.39
SMB	39.04	-16.62	2.99	11.52	21.99	2.17	0.26
HML	35.48	-13.45	4.50	12.38	15.63	1.84	0.36
WML	26.18	-78.96	14.46	27.53	18.24	-2.47	0.53

Table 3.1: The long-run performance of momentum compared to the Fama-French factors. 'KURT' stands for excess kurtosis and 'SR' for Sharpe ratio. The sample returns are from 1927:03 to 2011:12.

	α	t-stat	ρ	t-stat	R^2	OOS R^2	$\overline{\sigma}$	σ_σ
	Panel A: 1927:03 to 2011:12							
RMRF	0.0010	6.86	0.60	23.92	36.03	38.81	12.81	7.82
WML	0.0012	5.21	0.70	31.31	49.10	57.82	15.03	12.26
	Panel B: 1963:07 to 2011:12							
RMRF	0.0009	5.65	0.58	17.10	33.55	25.46	13.76	8.48
SMB	0.0004	8.01	0.33	8.32	10.68	-8.41	7.36	3.87
HML	0.0001	4.88	0.73	25.84	53.55	53.37	6.68	4.29
WML	0.0009	3.00	0.77	29.29	59.71	55.26	16.40	13.77

Table 3.2: AR (1) of 1-month realized variances. The realized variances are the sum of squared daily returns in each month. The AR (1) regresses each realized variance on its lag and a constant. The OOS R-squared uses the first 240 months to run an initial regression so producing an OOS forecast. Then uses an expanding window of observations till the end of the sample. In panel A the sample period is from 1927:03 to 2011:12. In panel B we repeat the regressions for RMRF and WML and add the same information for the HML and SMB. The last two columns show, respectively, the average realized volatility and its standard deviation.

	Max	Min	Mean	STD	KURT	SKEW	SR
WML	26.18	-78.96	14.46	27.53	18.24	-2.47	0.53
WML^*	21.95	-28.40	16.50	16.95	2.68	-0.42	0.97

Table 3.3: Momentum and the economic gains from scaling. The first row presents as a benchmark the economic performance of plain momentum from 1927:03 to 2011:12. The second row presents the performance of risk-managed momentum. The risk-managed momentum uses the realized variance in the previous 6 months.

	$CE(\mu_1)$	$CE(\mu_2)$	$CE(\mu_{i>2})$	$CE(r)$
RM	11.72	-7.39	-4.18	0.14
RM+WML	28.51	-13.90	-20.07	-5.46
RM+WML*	31.64	-13.00	-5.11	13.54

Table 3.4: The economic performance of momentum for a representative investor. The first row shows the performance of the market portfolio. The second row combines the market portfolio with momentum and the third one with scaled momentum. The first column shows the certainty equivalent of the mean return of each strategy. The second and third columns present the contribution to the certainty equivalent of standard deviation and higher moments, respectively. The last column shows the certainty equivalent obtained from annual non-overlapping returns. The returns are from 1927:03 to 2011:12. The decomposition uses a Taylor expansion of the utility function around the mean return of the portfolio with a CRRA of 4.

	α	t-stat	ρ	t-stat	R^2	R^2_{OOS}
σ^2_{wml}	0.0012	2.59	0.70	12.58	0.49	0.44
σ^2_{rmrf}	0.0012	4.29	0.50	7.37	0.25	0.07
β^2	0.3544	6.05	0.21	2.83	0.05	0.05
$\beta^2\sigma^2_{rmrf}$	0.0007	2.73	0.47	6.80	0.22	0.21
σ^2_{ε}	0.0007	2.69	0.72	13.51	0.52	0.47

Table 3.5: Risk decomposition of momentum risk. Each row shows the results of an AR (1) for 3-month, non-overlapping periods. The first row is for the realized variance of the WML and the second one the realized variance of the market. The third row is squared beta, estimated as a simple regression of 63 daily returns of the WML on RMRF. The fourth row is the systematic component of momentum risk and the last row the specific component. The OOS R-squares use an expanding window of observations after an initial in-sample period of 20 years.

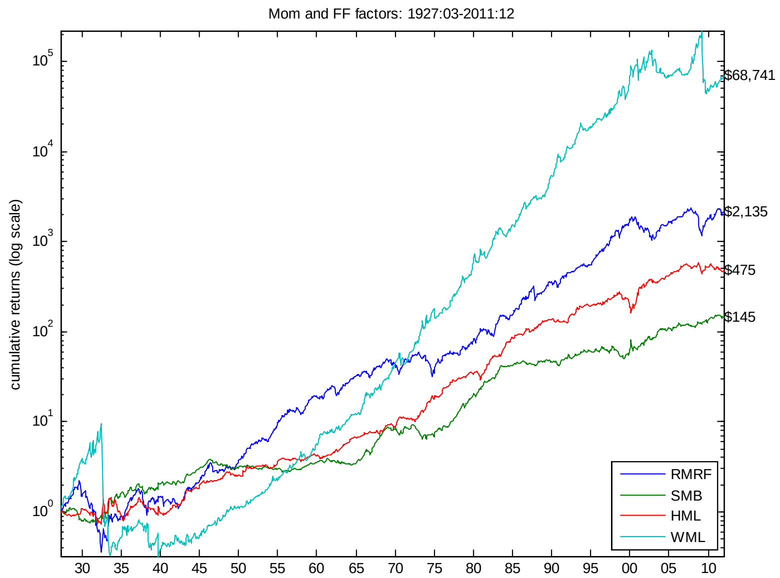

Figure 3.1: The long-run cumulative returns of momentum compared to the Fama-French factors. Each strategy consists on investing $1 at the beginning of the sample in the risk-free rate and combine it with the respective long-short portfolio. The proceeds are fully reinvested till the end of the sample. On the right is the terminal value of each strategy.

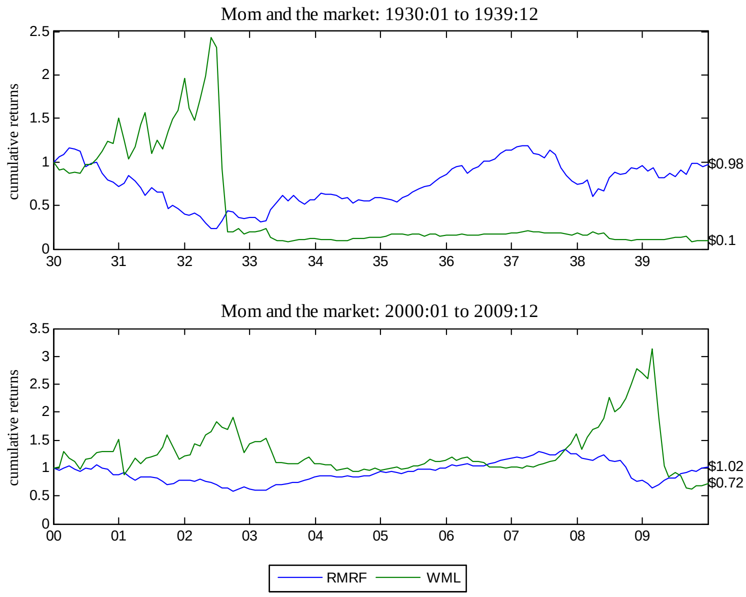

Figure 3.2: Momentum crashes. The figure plots the cumulative return and terminal value of the momentum and market portfolio strategies in its two most turbulent periods: the 1930's and the 2000's.

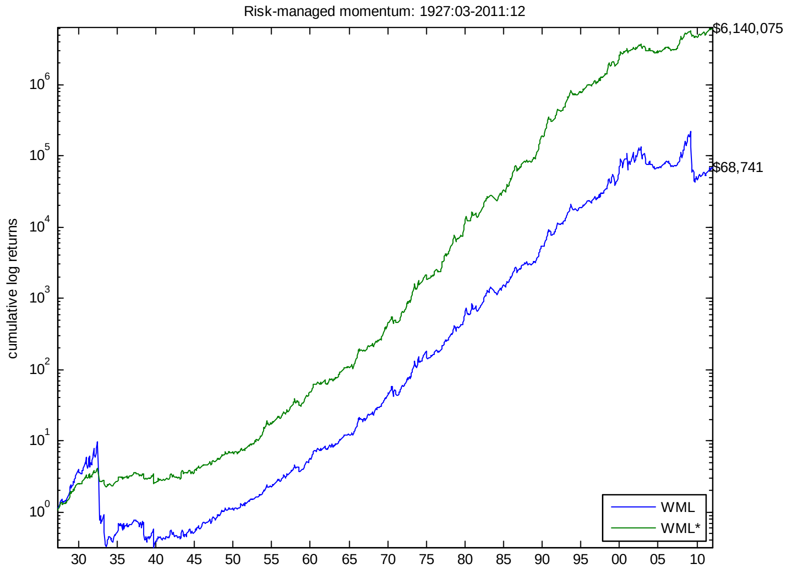

Figure 3.3: The long-run performance of risk-managed momentum. The risk-managed momentum (WML*) scales the exposure to momentum using the realized variance in the previous 6-months. In the beginning of the sample the strategy invests $1 in the risk-free asset and combines it with the long-short portfolio. The proceeds are fully reinvested till the end of the sample. On the right is the terminal value of the strategy.

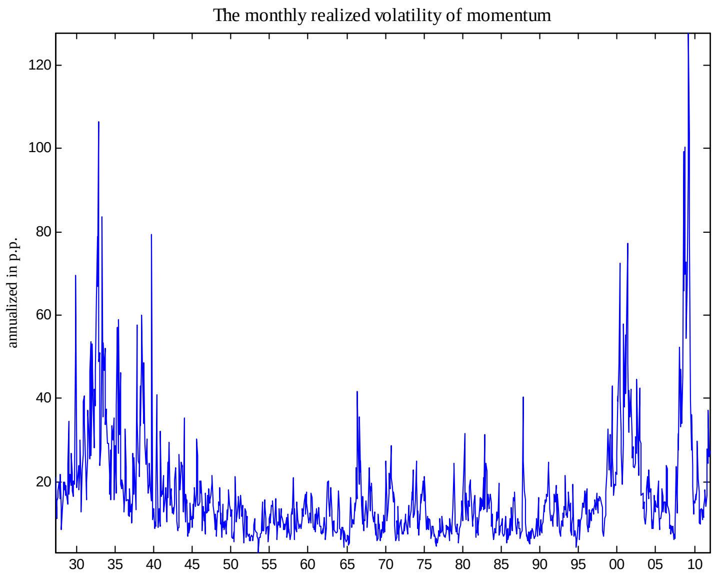

Figure 3.4: The realized volatility of momentum obtained from daily returns in each month from 1927:03 to 2011:12.

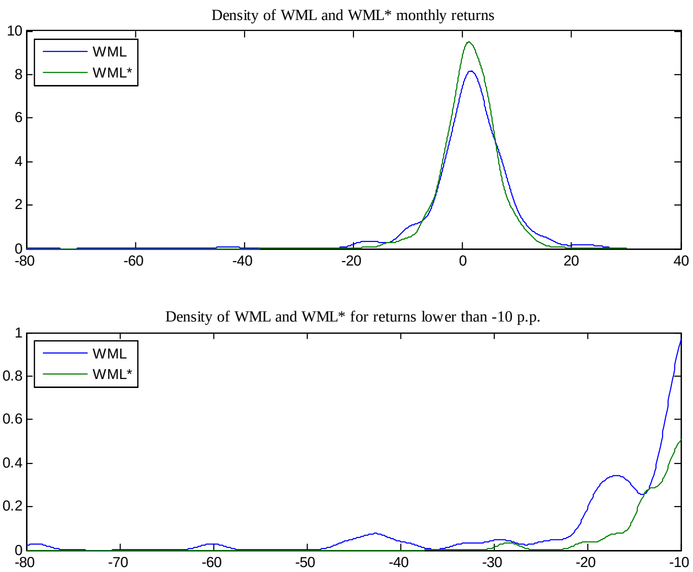

Figure 3.5: The density of plain momentum (WML) and risk-managed momentum (WML*). The risk-managed momentum uses the realized variance in the previous 6 months to scale the exposure to momentum. The returns are from 1927:03 to 2011:12.

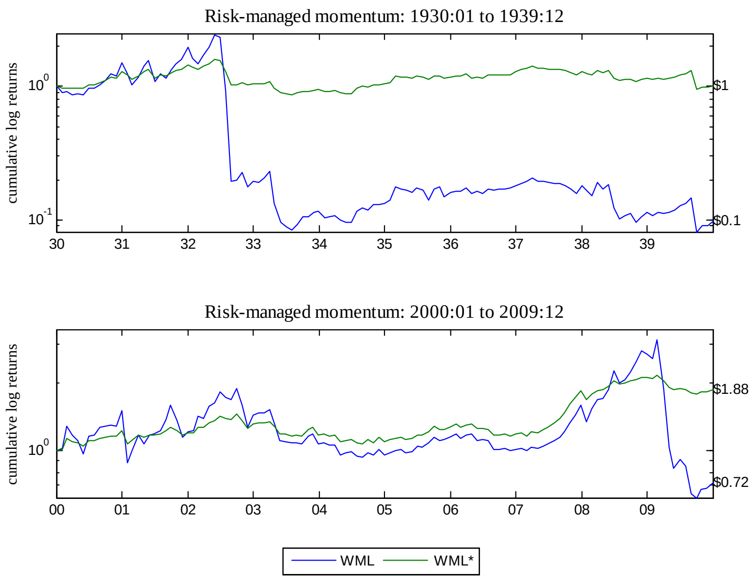

Figure 3.6: The benefits of risk-management in the 1930's and the 00's. The risk-managed momentum (WML*) uses the realized variance in the previous 6 months to scale the exposure to momentum.

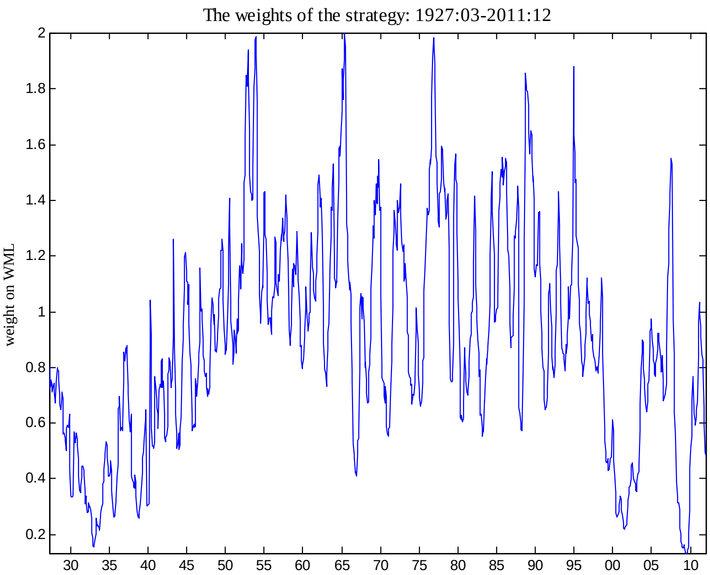

Figure 3.7: Weights of the scaled momentum. The risk-managed momentum uses the realized variance in the previous 6 months to scale the exposure to momentum.

Bibliography

[1] Ang, A., and J. S. Chen (2010). "Yield Curve Predictors of Foreign Exchange Returns." AFA 2011 Denver Meetings Paper. Available at SSRN: http://ssrn.com/abstract=1542342 .

[2] Asness, C. S., J. M. Liew, and R. L. Stevens (1997). "Parallels between the cross-sectional predictability of stock and country returns", Journal of Portfolio Management 23, pp. 79-87.

[3] Asness, C. S., T. J. Moskowitz, and L. H. Pedersen (2009). "Value and Momentum Everywhere." Working Paper, University of Chicago, Booth School of Business.

[4] Barro, R. (2006). "Rare Disasters and Asset Markets in the Twentieth Century." Quarterly Journal of Economics, 121, 823–866.

[5] Blitz, D., J. Huij, and M. Martens (2011). "Residual Momentum", Journal of Empirical Finance, Vol 18, 3. pp. 506-521.

[6] Berge, T., O. Jordà, and A. Taylor (2010). "Currency Carry Trades." in: NBER International Seminar on Macroeconomics, University of Chicago Press

[7] Bliss, R., and N. Panigirzoglou (2004). "Option-Implied Risk Aversion Estimates." Journal of Finance, Vol 59, 1, 407-446.

[8] Bollerslev, T. (1987), "A Conditional Heteroskedastic Time Series Model for Speculative Prices and Rates of Return", Review of Economics and Statistics, 69, pp. 542-547.

[9] Brandt, M. W., P. Santa-Clara, and Valkanov (2009). "Parametric Portfolio Policies: Exploiting Characteristics in the Cross-Section of Equity Returns." Review of Financial Studies, 22(9): 3411–3447.

[10] Brunnermeier, M. K., S. Nagel, and L. H. Pedersen (2008). "Carry Trades and Currency Crashes." NBER Working Paper no. 14473.

[11] Burnside, Craig (2011). "Carry Trades and Risk." NBER Working Paper No. 17278.

[12] Burnside, C., M. Eichenbaum, and S. Rebelo (2008). "Carry Trade: The Gains of Diversification." Journal of the European Economic Association 6(2-3), 581-588.

[13] Burnside, C., M. Eichenbaum, and S. Rebelo (2011). "Carry Trade and Momentum in Currency Markets." Annual Review of Financial Economics, forthcoming

[14] Burnside, C., M. Eichenbaum, I. Kleshchelski, and S. Rebelo. (2011). "Do Peso Problems Explain the Returns to the Carry Trade?" Review of Financial Studies, forthcoming.

[15] Carhart, M. M. (1997). "On Persistence in Mutual Fund Performance." Journal of Finance, 52, 57–82.

[16] Cheung, Y.-W., M. D. Chinn, and Pascual (2005). "Empirical exchange rate models of the nineties: Are any fit to survive?" Journal of International Money and Finance, 24(7), 1150–1175.

[17] Chinn, M. D. and H. Ito (2006). "What Matters for Financial Development? Capital Controls, Institutions, and Interactions." Journal of Development Economics 81, 163–192.

[18] Christiansen, C., A. Ranaldo, and P. Söderllind (2010). "The Time-Varying Systematic Risk of Carry Trade Strategies." Journal of Financial and Quantitative Analysis, forthcoming.

[19] Cochrane, J.H. (2011). "Presidential Address: Discount Rates." Journal of Finance, 66, 1047-1108.

[20] Daniel, K., and T. Moskowitz (2011). "Momentum Crashes", Working paper.

[21] Della Corte, P., L. Sarno, and I. Tsiakas (2009). "An Economic Evaluation of Empirical Exchange Rate Models." Review of Financial Studies 22, 3491-3530

[22] DeMiguel, V., L. Garlappi, and R. Uppal (2009). "Optimal Versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy?" Review of Financial Studies, 22(5), 1915-1953.

[23] Dornbusch, R. and S. Fischer (1980). "Exchange Rates and the Current Account." The American Economic Review, Vol. 70, N^{o}5, 960-971.

[24] Engle, R. (1982). "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of U.K. Inflation", Econometrica 50, pp. 987-1008.

[25] Engle, R., and T. Bollerslev (1986). "Modeling the Persistence of Conditional Variances", Econometric Reviews, 5, pp. 1-50.

[26] Erb, C. B., and C. R. Harvey (2006). "The strategic and tactical value of commodity futures", Financial Analysts Journal 62, pp. 69-97.

[27] Fama, E. F. (1984). “Forward and Spot Exchange Rates.” Journal of Monetary Economics, 14 (3), 319–338.

[28] Fama, E. F., and K. R. French (1992). “The Cross-Section of Expected Stock Returns.” Journal of Finance, 47, 427–465.

[29] Fama, E. and K. French (1996). “Multifactor Explanations of Asset Pricing Anomalies”, The Journal of Finance, vol. 51, no. 1, pp. 55-84.

[30] Farhi, E., and X. Gabaix (2007). “Rare Disasters and Exchange Rates: A Theory of the Forward Premium Puzzle.” Working Paper, Harvard University.

[31] Fung, W., and D. Hsieh (2004). “Hedge fund benchmarks: a risk based approach.” Financial Analyst Journal 60, 65–80.

[32] Gourinchas, P.-O. and H. Rey (2007). “International Financial Adjustment.” Journal of Political Economy, 115(4), 665–703.

[33] Gourio, F., M. Siemer, and A. Verdelhan (2011). “International Risk Cycles.” NBER Working Paper No. 17277.

[34] Goyal, A., and I. Welch (2008).“A Comprehensive Look at the Empirical Performance of Equity Premium Prediction”, Review of Financial Studies, 21 (4), 1455-1508.

[35] Grinblatt, M., and S. Titman (1989). “Mutual fund performance: An analysis of quarterly portfolio holdings”, Journal of Business, 62, pp. 394-415

[36] Grinblatt, M., and S. Titman (1993). “Performance measurement without benchmarks: An examination of mutual fund returns”, Journal of Business, 66, pp. 47-68.

[37] Grundy, B., and J. S. Martin (2001). “Understanding the nature of the risks and the source of the rewards to momentum investing”, Review of Financial Studies, 14, pp. 29-78.

[38] Jacobs, H., S. Müller, and M. Weber (2010). “How Should Private Investors Diversify? An Empirical Evaluation of Alternative Asset Allocation Policies to Construct a 'World Market Portfolio'.” Available at SSRN: http://ssrn.com/abstract=1471955 or http://dx.doi.org/10.2139/ssrn.1471955

[39] Jegadeesh, N. and S. Titman (1993). “Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency.” Journal of Finance, 48 (1), 65–91

[40] Jordà, O., and A. Taylor (2009). “The Carry Trade and Fundamentals: Nothing to Fear but FEER Itself.” NBER Working Paper No. 15518.

[41] Jordà, O., and A. Taylor (2011). "Performance of Zero Net-Investment Strategies." NBER working paper 17150.

[42] Jurek, J. W. (2009). "Crash-Neutral Currency Carry Trades"AFA 2010 Atlanta Meetings Paper. Available at SSRN: http://ssrn.com/abstract=1262934 or http://dx.doi.org/10.2139/ssrn.1262934

[43] Jylhä, P., and M. Suominen (2011). "Speculative Capital and Currency Carry Trades." Journal of Financial Economics, 99(1), 60-75.

[44] Kroencke, T., F. Schindler, and A. Schrimpf (2011). "International Diversification Benefits with Foreign Exchange Investment Styles." CREATES Research Paper No. 2011-10. Available at SSRN: http://ssrn.com/abstract=1790764 .

[45] Lesmond, D. A., M. J. Schill, and C. Zhou (2004). "The Illusory Nature of Momentum Profits." Journal of Financial Economics, 71(2), 349–380.

[46] Levich, R., and M. Pojarliev (2011). "Detecting Crowded Trades in Currency Funds." Financial Analysts Journal, Volume 67, N^{o}1.,26-39.

[47] Lustig, H. N., N. L. Roussanov, and A. Verdelhan (2011a). "Common Risk Factors in Currency Markets." Review of Financial Studies, forthcoming.

[48] Lustig, H. N., N. L. Roussanov, and A Verdelhan. (2011b). "The Cross-Section of Foreign Currency Risk Premia and Consumption Growth Risk: A Reply." American Economic Review, forthcoming.

[49] Lustig, H. N., A. Verdelhan (2007). "The Cross Section of Foreign Currency Risk Premia and Consumption Growth Risk." American Economic Review, 97, 89–117

[50] Mancini, L., A. Ranaldo, and J. Wrampelmeyer (2011). "Liquidity in the Foreign Exchange Market: Measurement, Commonality, and Risk Premiums." Journal of Finance, forthcoming.

[51] Meese, R. A. and K. Rogoff (1983). "Empirical Exchange Rate Models of the Seventies : Do They Fit Out of Sample?" Journal of International Economics, 14(1-2), 3-24.

[52] Mehra, R. and E.C. Prescott (1985). "The Equity Premium: A Puzzle." Journal of Monetary Economics, 15(2), 145–61.

[53] Melvin, M. and M. P. Taylor (2009). "The Crisis in the Foreign Exchange Market." CESifo Working Paper No. 2707.

[54] Menkhoff, L., L. Sarno, M. Schmeling, and A. Schrimpf (2011a). "Carry Trades and Global Foreign Exchange Volatility." Journal of Finance, forthcoming.

[55] Menkhoff, L., L. Sarno, M. Schmeling, and A. Schrimpf (2011b). "Currency Momentum Strategies." Available at SSRN: http://ssrn.com/abstract=1809776 .

[56] Menkhoff, L., L. Sarno, M. Schmeling, and A. Schrimpf (2012). "Currency Momentum Strategies", Journal of Financial Economics, forthcoming

[57] Moskowitz, T., and M. Grinblatt (1999). "Do Industries Explain Momentum?", Journal of Finance, vol. 54, no. 4, pp. 1249-1290.

[58] Obstfeld, M. and K. Rogoff (2005). "Global Current Account Imbalances and Exchange Rate Adjustments." Brookings papers on Economic Activity, N^{o}1, 67-123.

[59] Okunev, J. and D. White (2003). "Do Momentum-Based Strategies Still Work in Foreign Currency Markets?" Journal of Financial and Quantitative Analysis, 38(2), 425—447.

[60] Rapach, D.E., and M.E. Wohar (2002). "Testing the Monetary Model of Exchange Rate Determination: New Evidence from a Cen-

tury of Data." Journal of International Economics, 58, 359–385.

[61] Rogoff, K. (2009). "Exchange Rates in the Modern Floating Era: What Do We Really Know?" Review of World Economics, 145(1), 1–12.

[62] Rogoff, K. S., and V. Stavrakeva (2008). "The Continuing Puzzle of Short Horizon Exchange Rate Forecasting." SSRN eLibrary.

[63] Rouwenhorst, K. (1998). "International momentum strategies", Journal of Finance, 53, pp. 267-284.

[64] Rouwenhorst, K. (1999). "Local return factors and turnover in emerging stock markets", Journal of Finance 54, pp. 1439-1464.

[65] Schwert, G. (1989). "Why Does Stock Market Volatility Change over Time?", Journal of Finance, 44, pp. 1115-1154.