Tim Alexander Kroencke

Understanding and Harvesting Expected Returns of Asset Classes, Investment Styles, and Risk Factors

Inauguraldissertation

zur Erlangung des akademischen Grades

eines Doktors der Wirtschaftswissenschaften

der Universität Mannheim

Vorgelegt im Frühjahrssemester 2013

Dekan: Dr. Jürgen M. Schneider

Berichterstatter: Prof. Dr. Erik Theissen

Prof. Dr. Stefan Ruenzi

Tag der mündlichen Prüfung: 28. Mai 2013

Understanding and Harvesting Expected Returns of Asset Classes, Investment Styles, and Risk Factors

Thesis Advisor: Professor Dr. Erik Theissen Author: Tim Alexander Kroencke

Introduction

The Tale of Ketchup Economics. $P_t = E_t\left[M_{t+1}X_{t+1}\right]$, or the price of an asset today should be equal to the future expected discounted value of its payoffs given the marginal utility of the investor as the discount factor. This simple concept is central in asset pricing. However, there are two polar approaches of asset pricing. Both can be clarified using Larry Summers' famous tale of "ketchup economics", summarized as follows (Summers (1985)):

> There are two groups of economists studying the ketchup market. The first group is called general economists and the second group is called financial economists. General ketchup economists try to understand the prices of bottles of ketchup by looking at the market of tomatoes. They try to understand the fundamental sources of macroeconomic risk to determine the prices of bottles of ketchup. They are addicted to deep models with utility and production functions.
>
> Financial ketchup economists try to understand the prices of bottles of ketchup by looking at the ketchup market more deeply. They try to understand prices of bottles of ketchup relative to the prices of other bottles of ketchup. They love to exploit powerful arbitrage concepts.

In this thesis, I study investment returns using both perspectives. I deal with the "financial ketchup economics" perspective in chapter 3 and chapter 4. The central asset pricing equation is tested for a set of interesting test assets, whereby I approximate investor's marginal utility with the returns of some benchmark assets. This approach allows me to quantify if the test assets provide economically sizeable and significant returns which are not feasible with the benchmark assets (i.e. diversification benefits) and to learn how investors can efficiently harvest them.

Importantly, the bottom line of Summers (1985) is that there is insufficient research in the interstices between the financial and the economics approaches:

> "It is unfortunate [...] that researchers in economics pay so little attention to finance research, and perhaps more unfortunate that financial economists remain so reluctant to accept any research relating asset prices and fundamental values."

Thus, after studying how to harvest returns, I turn to understanding returns. In chapter 1 and chapter 2, I investigate which macroeconomic sources of risk can explain average investment returns.[1] In particular, these chapters are devoted to the construction and measurement of the fundamental risk factors driving expected returns of asset classes as well as popular investment styles. To this end, I follow Larry Summers' suggestion and make use of concepts and methods from both approaches of "ketchup economics".

Harvesting Returns. I analyze how investors can gain some extra return from their international investment portfolio. International diversification is intended to improve the return per unit of risk, in the hope that returns across countries are not perfectly correlated (e.g. Grubel (1968) and Solnik (1974)).

The chapter "*International Diversification with Securitized Real Estate and the Veiling Glare from Currency Risk*" shows that the risk and return characteristics of global stocks and bonds are different from those of global real estate. Due to its local nature, real estate is subject to rather local factors and less so to global factors (Eichholtz (1996)). Thus, diversification benefits from international (listed) real estate are potentially larger than the benefits from common stocks.

[1] I present the papers of my thesis in reverse chronological order.

After accounting for systematic FX risk (Lustig and Verdelhan (2007), Verdelhan (2011)), which adds inter-asset and inter-country correlation, the paper finds evidence for this hypothesis.

The recent literature has documented that particularly style-based investment strategies in FX markets provide attractive returns (e.g. Lustig and Verdelhan (2007), Pojarliev and Levich (2008), Menkhoff, Sarno, Schmeling, and Schrimpf (2012b), Asness, Moskowitz, and Pedersen (2012)). Thus, the FX component may not only add risk to a portfolio but can also be a source of additional diversification benefits.

The chapter "*International Diversification Benefits with Foreign Exchange Investment Styles*" studies the currency component of international stock and bond portfolios. While style-based investments and their role for portfolio allocation in equity markets have been widely studied (e.g. Eun, Huang, and Lai (2008), or Eun, Lai, de Roon, and Zhang (2010)), there is considerably less knowledge about the portfolio implications of style investing in FX markets. The paper quantifies economically large and significant diversification benefits from simple FX investment styles like carry, momentum and value. Almost all previous research focuses on currency "hedging" in a portfolio context (e.g. Glen and Jorion (1993), Campbell, de Medeiros, and Viceira (2010)). In contrast, this chapter documents that style-based currency "speculation" can be quite attractive in a global portfolio.

Understanding Returns. The remaining two chapters are located in the interstices between the two fields of "ketchup economics" and study the relationship between the real economy and investment returns. Before Summers (1985) as well as in the subsequent years, this kind of literature proceeded rather slowly with classic contributions by Chen, Roll, and Ross (1986), Breeden, Gibbons, and Litzenberger (1989), and Campbell (1996). The main issue is that the correlation between standard macroeconomic variables and financial returns is low, and thus, the returns of risky asset classes like equity (Mehra and Prescott (1985)), or specific investment styles like value and momentum (Fama and French (1993), Jegadeesh and Titman (1993)), seem to be too large to be justifiable.

More recently, starting with the new millennium, more promising attempts have been made by introducing novel macroeconomic variables as fundamental risk factors which are more successful (to name a few contributions, Lettau and Ludvigson (2001), Vassalou (2003), Parker and Julliard (2005), Petkova (2006), Jagannathan and Wang (2007), Savov (2011), Koijen, Lustig, and Nieuwerburgh (2012)). Indeed, today there is a "zoo" of fundamental factors which successfully explain several kinds of investment returns (Cochrane (2008)). As a consequence, it is quite difficult to assess how all these variables relate to each other. More specifically, many of these fundamental factors produce mixed results in different empirical settings, or differently motivated factors lead to very similar results. Thus, with these two chapters of my thesis at hand, I aim to analyze in a systematic manner properties, construction and measurement of several newly introduced fundamental risk factors to gain novel insights.

The chapter "*GDP Mimicking Portfolios and the Cross-Section of Stock Returns*" studies popular GDP-based macroeconomic factors in a unified framework. Aggregate GDP as a measure of fundamental risk is a popular risk factor in asset pricing. From an empirical perspective, aggregate GDP does not correlate well with stock returns and is not useful for explaining average returns. However, there is a stylized fact in macroeconomics, which has been so far ignored in finance: some components of GDP lead the aggregate, while some components of GDP lag the aggregate (e.g. Greenwood and Hercowitz (1991), Gomme, Kydland, and Rupert (2001), Davis and Heathcote (2005), Fisher (2007), Leamer (2007)). The paper documents that leading GDP components can explain the size premium and the value premium quite well. The opposite is documented for the momentum premium. The lagging GDP components explain the return of momentum portfolios very well. A three-factor model with the market excess return, one leading and one lagging GDP component compares very favorably with the Carhart four-factor model in jointly explaining a large cross-section of size, book-to-market, momentum, and industry portfolio returns.

Finally, the chapter "*Asset Pricing without Garbage*", which also serves as my Job Market Paper, looks more deeply at consumption as a fundamental risk factor. The key contribution is that the paper exploits insights from state space analysis to provide an explanation for the bad performance of traditional consumption measures (Mehra and Prescott (1985), Breeden, Gibbons, and Litzenberger (1989)) and the good performance of recently proposed alternative consumption measures in the classical consumption-based asset pricing model (Parker and Julliard (2005), Jagannathan and Wang (2007), Savov (2011)). More specifically, the paper provides an explanation of why garbage as a measure of consumption implies a several times lower coefficient of relative risk aversion in the consumption-based asset pricing model than consumption based on the official National Income and Product Accounts (NIPA). Unlike garbage, NIPA consumption is filtered to mitigate measurement error. The paper applies a structural model of the filtering process, which allows to revoke the filter inherent in NIPA consumption. "Unfiltered NIPA consumption" performs as well as garbage in explaining the equity premium and risk-free rate puzzle. Furthermore, the paper documents that two other popular NIPA-based measures, three-year and fourth-quarter NIPA consumption, are related to unfiltered NIPA consumption. Both can be viewed as ad hoc unfilter rules.

Contents

List of Tables

List of Figures

Acknowledgements

The origin of my interest for different asset classes, investment styles and risk factors goes back to my undergraduate studies at University of Freiburg and University of Basel. I was inspired by Heinz Rehkugler to study different asset classes (stocks, real estate, foreign exchange) as well as different investment styles (value, momentum, carry). Heinz Zimmermann and Yvan Lengwiler taught me to apply one equation, $P_t = E_t\left[M_{t+1}X_{t+1}\right]$, to study asset pricing from the two perspectives of "ketchup economics". What I have learned from them shaped my research.

After graduation, I got the opportunity to conduct my own research at ZEW Mannheim and University of Mannheim. During my first projects (chapters 2 to 4), I was taught step by step to carry out a research project from beginning to end by my advisor Erik Theissen at University Mannheim and by my former ZEW colleagues Felix Schindler and Andreas Schrimpf. This was in no way as easy as it sounds. However, all three encouraged and motivated me through out this process such that it was a great experience that I never want to miss. I am immensely grateful for their guidance and advice. The first chapter of my thesis, which is also my first single authored paper, is the ultimate synthesis of what I have learned during this time.

I also owe thanks to Wolfgang Franz, Michael Schroeder, and Peter Westerheide from ZEW Mannheim, for giving me the opportunity to write and fund my thesis at ZEW, Stefan Ruenzi from University of Mannheim, who encouraged me to affiliate with University of Mannheim, as well as Steffen Sebastian from University of Regensburg, who co-authored one chapter of my thesis.

I also want to thank others who have contributed to parts of my thesis. Chapter 1 has benefited from comments by John Cochrane, Lena Jaroszek, Frieder Mokinski, Jonathan Parker and participants at the University of Mannheim Financial Markets Seminar, and the ZEW Brownbag. I am grateful to Alexi Savov for sharing his data with me.

Chapter 2 has benefited from comments by Andre Betzer, Roland Fuess, Marco Pagano, Bertram Steininger, and participants of the 2012 annual meeting of the German Finance Association's Annual Meeting and seminar participants at the universities of Marburg, Giessen, Konstanz, Naples, Singapore, and St. Gallen.

Chapter 3 has benefited from comments by Keith Brown, Ines Chaieb, Charlotte Christiansen, Anders Ekholm, Roland Fuess, Stephan Jank, Richard Levich, Lukas Menkhoff, Jesper Rangvid, Maik Schmeling, Stephan Siegel, Philippe Versijp, Paolo Vitale, Pim van Vliet and participants at the McGill Conference on Global Asset Management 2011, the CFR Asset Management Colloquium 2011, the Humboldt-Copenhagen Conference on Financial Econometrics 2011, the FMA European Conference in Porto 2011, the German Finance Association's Annual Meeting 2011, the Australasian Finance and Banking Conference in Sydney 2011, the Tor Vergata Conference on Banking and Finance in Rome 2011, the German Economic Association's Annual Meeting 2012, and seminar participants at Deka Investment. I gratefully acknowledge the BlackRock Research Award for the best paper on capital markets, fund management, mutual funds at the 2011 Australasian Finance and Banking Conference.

Chapter 4 has benefited from comments by Jim Clayton, Roland Fuess, Paskalis Glabadanidis, Seow Eng Ong, Peter Westerheide, Chih-Yuan Yang, and participants at the Tilburg Conference on International Real Estate Securities 2011, the Asia-Pacific Real Estate Research Symposium 2011, the 27th Annual Conference of the American Real Estate Society, the 17th Annual Conference of the Pacific-Rim Real Estate Society, the Annual Meeting of the German Finance Association 2011, and the Annual Meeting of the German Economic Association 2011. I gratefully acknowledge the award for the best paper presented at the Conference of the Pacific Rim Real Estate Society 2011, and financial support from the Leibniz Network on Real Estate Markets and Capital Markets (ReCapNet).

My final thanks are to my remaining ZEW colleagues and to all others from the Graduate School of Mannheim Universitiy, whom I spent a lot of time with during the four years of writing my thesis. It was great to philosophize with them about finance as well as Bundesliga.

February 28, 2013

To Veronika.

Chapter 1

1 Asset Pricing without Garbage

This chapter is single authored and my Job Market Paper. I present the papers of my thesis in reverse chronological order.

1.1 Introduction

The seminal paper by Savov (2011) finds that the consumption-based asset pricing model (CCAPM) using garbage as a measure of consumption matches the equity premium and risk-free rate with a several times lower coefficient of relative risk aversion than any other consumption measure based on the National Income and Product Accounts (NIPA) provided by the U.S. Bureau of Economic Analysis (BEA). Reported NIPA consumption growth has a lower sample standard deviation, a several times higher autocorrelation, and a significantly lower stock market covariance than garbage growth.[2] A possible explanation for the relative success of garbage is that reported NIPA consumption fails to measure consumption properly (Savov (2011, p. 200)). However, one has to keep in mind that dozens of brainy statisticians make a considerable effort to estimate NIPA consumption as precisely as possible. Thus, it is highly unlikely that NIPA consumption is simply badly measured and that the story ends here.

Furthermore, two recent papers report a surprising success of specific versions of NIPA consumption in explaining the cross-section of 25 Fama-French portfolio returns (Fama and French (1993)). The first is "ultimate consumption risk" by Parker and Julliard (2005), which is measured as three-year growth of NIPA consumption. The second is fourth-quarter to fourth-quarter growth of NIPA consumption proposed by Jagannathan and Wang (2007). Both NIPA-based alternative consumption measures are priced in the cross-section of stock returns, and there is evidence that they produce lower average pricing errors in cross-sectional regressions than

[2] The terms "consumption" and "consumption growth" are used interchangeably throughout the paper unless indicated otherwise.

garbage (Savov (2011)). However, in contrast to garbage, the NIPA-based alternative consumption measures are unable to explain the equity premium in combination with the risk-free rate (Savov (2011)). Nevertheless, the cross-sectional success suggests that there is valuable information inherent in NIPA consumption. In summary, there are three alternative consumption measures: garbage, three-year NIPA consumption, and fourth-quarter to fourth-quarter NIPA consumption. Each of these measures is either able to explain one or the other aspect of the data. However, so far, no unifying explanation exists which relates all these measures to each other, and a better understanding of this "alternative consumption measures puzzle" is highly warranted.

This paper entertains the following explanation. Each component of officially measured consumption is filtered to account for measurement error. Since the estimation methodology follows a bottom-up principle, the filter is applied at the level of each component of consumption (e.g. food, clothing, health care). As a result, the filter of individual consumption components is carried over to the aggregate consumption measure (e.g. nondurables and services). However, for aggregate consumption, the optimal filter is likely to be much smaller. The simpler measure of garbage, on the other hand, is not subject to filtering, and thus not distorted in this way. Moreover, filtering is *intensified* by the well-known bias stemming from time aggregation. Reported consumption is an estimate of consumption flow during a specific period, however, the CCAPM relates asset returns to consumption at one specific point in time (Breeden, Gibbons, and Litzenberger (1989)). As shown in this paper, filtering intensified by time aggregation is disastrous for the stochastic properties of reported NIPA consumption and generally matches the empirically observed differences to garbage.

I apply a structural model to approximate the filtering process in the consumption parts of NIPA. The model allows to revoke the filter inherent in *aggregate* NIPA consumption in a simple closed form. I call the measure emerging from this process "unfiltered NIPA consumption". The hypothesis I can test is that the stochastic properties of unfiltered NIPA consumption are similar to those of garbage and mitigate the equity premium and risk-free rate puzzle.

I also find evidence that unfiltered NIPA consumption is related to three-year consumption (Parker and Julliard (2005)) and fourth-quarter to fourth-quarter consumption (Jagannathan and Wang (2007)). A simulation experiment shows that fourth-quarter to fourth-quarter consumption is effective in removing time aggregation. In contrast, the three-year measure is effective in removing filtering. Thus, both can be viewed as simple ad hoc unfilter rules. However, in comparison to the direct unfilter method, both measures are less eligible in removing time aggregation and filtering at the same time.

Unfiltered NIPA consumption performs well in both disciplines (equity premium and cross-section). First, I find that unfiltered NIPA consumption is able to explain the equity premium and the risk-free rate with a coefficient of relative risk aversion between 18 and 34 in the post-war period (1950 - 2011), which is close to a coefficient of 15 implied by garbage. A feature of unfiltered NIPA consumption is a sample covering a longer period. When including the pre-war observations in the sample (1927 - 2011), I find a relative risk aversion as low as 9 for unfiltered NIPA consumption. Importantly, unfiltered NIPA consumption can explain the equity premium and the risk-free rate in joint tests, as does garbage. Second, unfiltered NIPA consumption implies a priced consumption risk premium in the cross-section of the 25 Fama-French portfolio returns, and gives a cross-sectional fit similar to the NIPA-based alternative consumption measures proposed by Parker and Julliard (2005) and Jagannathan and Wang (2007). In the full sample, as well as in the shorter post-war sample, I find a cross-sectional R^2 exceeding one half and low average pricing errors for unfiltered NIPA consumption.

The unfilter approach applied in this paper was first introduced by Quan and Quigley (1989), Geltner (1989), and Geltner (1993) in a different context. Research in real estate finds that appraisal-based real estate indices have a lower standard deviation, a substantially larger autocorrelation, and a lower stock market covariance than market value-based real estate indices. Before the work of Quan, Quigley and Geltner, it was a widely accepted fact that appraisal-based values of real estate are simply prone to measurement error, and thus less reliable than market values. Quan, Quigley and Geltner shifted the focus of the literature on the appraisal process, and documented possible reasons for - and the effects of - filtering on time series of real estate indices. In this paper, I argue that "appraised" consumption is subject to very similar issues, and indeed, there is related research supporting this view.

Virtually every paper on consumption-based asset pricing mentions possible problems related to measurement imperfections of consumption. However, surprisingly little work is devoted to analyzing the influence of the measurement methodology on the properties of reported consumption time series. As a rare exception, Wilcox (1992) and Bell and Wilcox (1993) focus on the retail trade survey, which is the most important ingredient of NIPA nondurable consumption. Wilcox (1992) discusses several sources of measurement errors in the construction of U.S. consumption data induced by the retail trade survey, how these imperfections are treated, and summarizes implications for empirical work. Bell and Wilcox (1993) warn that conclusions "about underlying theoretical parameters may be sensitive to imperfections in the data", and propose empirical adjustments to take this issue into account.[3]

The retail trade survey is an important ingredient of NIPA consumption, however, the consumption estimates are subject to several additional sources and issues (BEA (2009)). The sources of NIPA consumption include the Census Bureau, the Bureau of Labor Statistics, the Energy Information Administration, the Internal Revenue Service, and the U.S. Department of Agriculture. These agencies conduct several surveys which take place at different frequencies. Already at this level, a "variety of methods" is used to extract consumption estimates from these surveys (BEA (2009)). For some items, no data is available at all and a residual method is used, i.e. private consumption is measured as residual to business and government purchases.[4] Following this, final estimates are benchmarked to more comprehensive but less frequently conducted surveys on an annual and quinquennial basis. The benchmarking procedure is described in the "NIPA handbook" (BEA (2009)). In brief, non-benchmark years are treated as indicator series

[3] Gregory and Wirjanto (1993) demonstrate how these imperfections may affect results from empirical tests of the CCAPM.

[4] As Savov (2011) points out, business and government purchases can be subject to interpolation or even judgement in place of data. The exact estimation procedure differs between the components of total consumption; the "NIPA handbook" of the BEA provides an overview (BEA (2009)).

which are used to interpolate and extrapolate expenditure patterns from benchmark years. Bell and Wilcox (1993, p. 264) note that benchmarking involves a filtering of the estimates, and thus, "will affect their autocorrelation properties".[5] The internet appendix to Savov (2011) provides a description of many of these issues and how they may account for differences to garbage.[6]

There is considerably more work focusing on the time aggregation bias only. For example, Hall (1988) and Grossman, Melino, and Shiller (1987) propose some variants of (un)filtering methods to account for time aggregation in the estimation method. Ait-Sahalia, Parker, and Yogo (2004) report estimates of the coefficient of relative risk aversion, which are simply corrected for time-aggregation by dividing by two. Also the influential paper by Bansal and Yaron (2004) carefully accounts for time aggregation when the model is calibrated to match empirical moments.

As mentioned before, two recent contributions propose the direct use of alternative NIPA-based consumption measures. Parker and Julliard (2005) argue that ultimate consumption risk can provide a "correct measure of risk under several extant explanations of slow consumption adjustment". As their first example for slow consumption adjustment, Parker and Julliard (2005, p. 186) mention "measurement error in consumption". Jagannathan and Wang (2007) consider that investors are more likely to simultaneously make consumption and investment decisions in the fourth-quarter of each year, when the investors' tax year ends, and that the end-of-year consumption is thus a better measure for asset pricing.[7] However, using fourth-quarter to fourth-quarter consumption is a straightforward way to mitigate time aggregation and to bring the data closer to point consumption growth as well (Breeden, Gibbons, and Litzenberger (1989), Savov (2011)).

[5] Indeed, Bell and Wilcox (1993) use data on the retail trade survey *before* benchmarking. I tried to acquire similar non-benchmarked retail trade data to better pin down the effect of benchmarking on the data. Although the Census Bureau has been very helpful with respect to questions related to the retail trade survey, it was not possible for me to obtain non-benchmarked data.

[6] I focus on annual data, however, when turning to higher frequencies, seasonal adjustments are a further layer of filtering. Ferson and Harvey (1992) study the CCAPM using seasonally unadjusted consumption data. They find that a model with seasonally unadjusted data and non-separable preferences with seasonal effects works better than the standard model.

[7] Jagannathan, Marakani, Takehara, and Wang (2012) provide some evidence for the infrequent portfolio allocation decision hypothesis using an international comparison where the end of investors' tax year falls on different quarters.

Breeden, Gibbons, and Litzenberger (1989) discuss the possible influence of interpolation *and* time aggregation in consumption data. They are particularly concerned that interpolation is "exacerbated" by time aggregation, and conclude that "it is difficult to disentangle the two effects" (Breeden, Gibbons, and Litzenberger (1989, p. 243)). Indeed, to my knowledge, their is so far no paper to cite which tried to disentangle both effects.

The next section (1.2) will present a structural model of the filter process inherent in NIPA data. The model allows to revoke the filter inherent in NIPA consumption using a simple formula. Furthermore, a small adjustment is suggested in order to take time aggregation into account. A simulation experiment disentangles the two effects and provides a comparison of the mechanics of three-year consumption and fourth-quarter to fourth-quarter consumption as ad hoc unfilter rules. Section 1.5 provides the empirical properties of the alternative consumption measures. I use these observations to calibrate "unfiltered NIPA consumption" to garbage. In a next step, I present estimates on the coefficient of relative risk aversion as well as on the consumption risk premium using cross-sectional regressions. Section 1.6 concludes.

1.2 Measurement Error, Filtering, and Time Aggregation in Consumption Data

It is not feasible to model all the different layers in the construction of NIPA consumption exactly. Some of the processes involved are not fully known and the original data is not available. However, a more parsimonious approach in the spirit of the adjustments of Bell and Wilcox (1993) or Hall (1988) is promising. I apply a structural model hoping to account for the most important aspects of filtering in the measurement methodology. The main benefit of this approach is a simple closed form solution for calculating "unfiltered NIPA consumption", which is easy to interpret and allows to disentangle additional effects from time aggregation. Whether the structural model is a good approximation of reality is an empirical question which will be addressed in Section 1.5.

1.3 Stylized Model

I assume that the state of $j = 1, ..., J$ components of "true" log consumption, $c_{j,t} = log\,(C_{j,t})$ follow a Gaussian random walk (Hall (1978)):

$$c_{j,t} = c_{j,t-1} + \sigma_{j,\eta}\eta_{j,t}, \tag{1}$$

where $\eta_{j,t} \sim N\,(0,\ 1)$ is economic disturbance. However, the true state of consumption is not observable. Instead, the observed time series $y_{j,t}$ satisfies

$$y_{j,t} = c_{j,t} + \sigma_{j,\xi}\xi_{j,t}, \tag{2}$$

where $\xi_{j,t} \sim N\,(0,\ 1)$ represents measurement error. Suppose, the goal of a statistician is to update the knowledge of the jth consumption component given the arrival of new information $y_{j,t}$. Let $\hat{c}_{j,t} = E_t\,(c_{j,t})$ be such an estimate of $c_{j,t}$ given the information set $F_t = \{y_{j,1}, ..., y_{j,t}\}$. I call $\hat{c}_{j,t}$ NIPA consumption in the following. Standard textbooks suggest a recursive projection to recover consumption given F_t (Ljungqvist and Sargent (2004), Tsay (2005)):[8]

$$\hat{c}_{j,t} = \hat{c}_{j,t-1} + \nu_j\,[y_{j,t} - \hat{c}_{j,t-1}]\,, \tag{3}$$

$$\nu_j = \frac{\sigma^2_{j,\eta}}{\sigma^2_{j,\xi} + \sigma^2_{j,\eta}}. \tag{4}$$

Equation (3) implies that the best estimate of NIPA consumption is the NIPA consumption estimate of the past period plus a weighted surprise. Intuitively, if measurement error is large relative to economic disturbance, ν_j will be small, and the expectation on the state of the jth component should be only slightly updated. In contrast, if there is no measurement error, it is optimal to adjust the expectation on the state of consumption one for one according to the observed variable.

[8] In essence, this is the Kalman (1960) filter for the local level model in state space analysis.

Equation (3) can be rearranged and expressed in terms of log growth rates:

$$\Delta\hat{c}_{j,t} = \nu_j \Delta y_{j,t} + (1 - \nu_j)\, \Delta\hat{c}_{j,t-1}, \tag{5}$$

The construction of aggregate consumption follows the bottom-up principle. First, each component is estimated separately and subsequently added up to total consumption. To keep the notation simple, the weighting is identical for all consumption components, i.e., $\nu = \nu_j$ for $j = 1, \ldots, J$. Therefore, the growth rate of *aggregate* NIPA consumption follows:

$$\Delta\hat{c}_t \cong \nu \Delta y_t + (1 - \nu)\, \Delta\hat{c}_{t-1}. \tag{6}$$

In Equation (6), the optimal filter ν for an individual component of consumption is carried over to aggregate consumption. This is a direct result of the bottom-up principle. The filtering reflects the need to estimate each consumption component separately and as accurately as possible. It is likely that the optimal level of filtering is much smaller for total consumption than for an individual component of consumption, and in terms of aggregate consumption, the individual filter is too strong. However, given the structural model, it is straightforward to recover an estimate for unfiltered NIPA consumption by solving Equation (6) backwards:

$$\Delta\hat{y}_t = \left[\Delta\hat{c}_t - (1 - \nu)\, \Delta\hat{c}_{t-1}\right] / \nu, \tag{7}$$

where $\Delta\hat{c}_t$ corresponds to aggregate NIPA consumption, and $\Delta\hat{y}_t$ is aggregate "unfiltered NIPA consumption".[9] Unfiltered NIPA consumption revokes the filter inherent in aggregate NIPA consumption and can be viewed as true consumption plus an additional error term which reflects (Geltner (1993)): First, the non-diversifiable part of the measurement error of the individual consumption components. Second, the deviations from the assumed (or calibrated) value ν to the true but unobservable filter. Third, all other simplifying assumptions of the structural

[9]This backwards induction was proposed by Geltner (1993) to estimate real estate market returns from appraisal-based real estate indices.

model. If the proposed model can account for important aspects of the data, unfiltered NIPA consumption should share the interesting properties of garbage, which can be thought of as observable consumption.

1.4 Accounting for Time Aggregation

Time aggregation is another layer of distortion when measuring consumption data and is more widely covered in the literature (e.g. Breeden, Gibbons, and Litzenberger (1989)) than filtering. The following suggests a simple modification of the unfilter rule in Equation (7) to account for time aggregation.

Consumption is usually measured as the flow of consumption during the interval of a specific period, and not as "spot" consumption at one point in time. However, it is spot consumption which conceptually enters the CCAPM. Working (1960) shows that if a variable follows a random walk, the variance of the growth rate of the time aggregated (logarithmic) variable is approximately 2/3 that of the true spot growth rate. Furthermore, the first-order autocorrelation is shifted from zero to 1/4, and, as shown by Taio (1972), the covariance to a second variable (e.g. the stock market) is reduced by 1/2.[10] Using a linearized version of the CCAPM, the coefficient of relative risk aversion is a function of the stock market covariance with consumption (see e.g. Campbell (2003)), and thus, time aggregation will bias estimates on the coefficient of relative risk aversion upwards by a factor of two (Breeden, Gibbons, and Litzenberger (1989)).

There are at least three ways to take time aggregation into account. The first, and most obvious, is to use December to December (or the fourth-quarter to fourth-quarter) consumption growth to reduce time aggregation in the first place (Breeden, Gibbons, and Litzenberger (1989), Savov (2011)). Unfortunately, this option is not available for consumption data in the prewar sample, where only annual consumption data is available. Second, a relative risk aversion estimate is simply adjusted by dividing by two, since the measured stock market covariance

[10]The approximation is exact for infinitesimal measurement intervals, as shown by Breeden, Gibbons, and Litzenberger (1989). In discrete time, these predictions are quite accurate for annual consumption time aggregated on monthly intervals, as can be inferred from Working (1960).

is only one-half of the true covariance (e.g. Breeden, Gibbons, and Litzenberger (1989), Ait-Sahalia, Parker, and Yogo (2004)). Third, Hall (1988) suggests to account for time aggregation using a simple autoregressive representation as an approximation.[11] In this spirit, Equation (7) can be applied as:

$$\Delta c_t^S = \left[\Delta c_t^{TA} - (1-\alpha)\,\Delta c_{t-1}^{TA}\right]/\alpha, \tag{8}$$

to adjust for time aggregation as well, where Δc_t^S is an estimate of spot consumption and Δc_t^{TA} is the time aggregated measure. It is easy to show that the parameter $\alpha = .80$ ensures that Δc_t^S has the same standard deviation as spot consumption.[12] Thus, an unfilter rule accounting for time aggregation *and* filtering can be set up as:

$$\Delta \hat{y}_t = \left[\Delta \hat{c}_t - (1-\phi)\,\Delta \hat{c}_{t-1}\right]/\phi, \tag{9}$$

where $\phi = \alpha \times \nu$. The time aggregation adjustment improves the stock market covariance of consumption through a larger standard deviation. The effect on the stock market correlation is not clear. To shed light on these effects, the following section examines how accurate the modified unfilter rule actually is.

1.4.1 Simulation Experiment

Intuitively, filtering destroys the correlation of consumption growth with stock returns, and drives a wedge between garbage and NIPA consumption. To quantify these effects, I simulate a model economy in which consumption-based asset pricing does work. Time aggregation and filtering

[11] More specifically, to apply the instrumental variables estimator by Hayashi and Sims (1983), Hall uses a second-order forward autoregressive (un)filter to remove autocorrelation of the time aggregated variables. Grossman, Melino, and Shiller (1987) suggest a maximum likelihood estimation strategy taking time aggregation directly into account.

[12] The value $\alpha = .80$ solves:

$$Var\left(\Delta c_t^{TA}\right) = \frac{\alpha}{2-\alpha} Var\left(\Delta c_t^S\right) = \frac{2}{3} Var\left(\Delta c_t^S\right),$$

where the first equality is implied by Equation (8) and the second equality uses the results in Working (1960) and Breeden, Gibbons, and Litzenberger (1989).

are added to true consumption to simulate a NIPA consumption measure. The performance of NIPA consumption, unfiltered NIPA consumption, long-run consumption (Parker and Julliard (2005)) and fourth-quarter consumption (Jagannathan and Wang (2007)) are compared with each other.

Asset Pricing Economy. The simulation is set up such that the pricing equation of the consumption-based model holds:

$$E\left[\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} R^e_{m,t+1}\right] = 0, \tag{10}$$

where β is the subjective discount factor, γ is the coefficient of relative risk aversion, and R^e_m is the simple calculated return of the market portfolio in excess of the risk-free rate (equity premium).

I generate 10,000 time-series of the model and take monthly observations of consumption growth and the equity premium from each run such that I can convert them to 60 annual observations. Simulated true consumption growth is calibrated to match with garbage, i.e. the annual standard deviation is 3% and the true correlation to the stock market is 0.60. Thus, I assume that measurement error for *aggregate* consumption is infinitesimally small and as a result observable consumption ("Garbage") is indistinguishable from true consumption. I set the true coefficient of relative risk aversion to 15. The resulting annual equity premium of this economy is 5.1% carrying a standard deviation of 20%.

In a second step, based on the monthly observations, I calculate a time aggregated annual "NIPA" consumption measure, and a second time aggregated and filtered NIPA consumption measure. The value of the filter parameter is set to $\nu = 1/2$.[13]

Finally, I calculate the "unfiltered NIPA consumption" measure, the P-J measure as in Parker and Julliard (2005), and the Q4Q4 as in Jagannathan and Wang (2007). For unfiltered NIPA consumption, several parameter values $\phi = \alpha \times \nu$ are applied. Taken from the discussion above,

[13] I find in Section 1.5 that a filter parameter ν of 1/2 is in line with empirical observations.

the recommended unfilter parameter is $\phi = .80 \times 1 = .80$ when NIPA consumption is only time aggregated, and $\phi = .80 \times 1/2 = .40$ when NIPA consumption is time aggregated and filtered. I also report results for further parameter values of ϕ as a sensitivity check. Further details on the simulation can be found in Appendix A.1.

Results. Table 1 provides characteristics of different consumption measures from 10,000 simulations of a consumption-based asset pricing economy. From left to the right, the table reports the properties of true consumption ("Garbage"), NIPA consumption, and alternative NIPA-based consumption measures using different unfilter rules. Panel A presents results for when NIPA consumption is only time aggregated and Panel B displays results for when NIPA consumption is time aggregated and filtered.

Table 1: Alternative Consumption Measures in a Simulated Asset Pricing Economy

The table displays statistics on different consumption growth measures in a consumption-based asset pricing economy, $E\left[\beta\left(C_{t+1}/C_t\right)^{-\gamma} R^e_{m,t+1}\right] = 0$. The economy is set up such that true consumption growth matches the empirical characteristics of garbage, i.e. a standard deviation of 3% and a correlation with stock returns of 0.60. The true coefficient of relative risk aversion (γ) is 15 which implies an equity premium of 5.1% (with a standard deviation of 20%). Displayed statistics are the median of 10,000 simulations of monthly data converted to an annual frequency with 60 observations. The reported coefficient of relative risk aversion (RRA γ) is a GMM estimate based on simulated consumption growth and the equity premium. Panel A shows the effect of time aggregation and Panel B shows the effect of time aggregation and filtering on the properties of consumption as reported in the NIPA. The columns to the right compare three different unfilter rules for the NIPA measure: The first two are three-year consumption growth (P-J) as in Parker and Julliard (2005), and fourth-quarter to fourth-quarter consumption growth (Q4Q4) as in Jagannathan and Wang (2007). The third is the new consumption measure using the direct unfilter method proposed in the text. Further simulation details are provided in the Appendix.

Panel A	Time Aggregation						
	Consumption Growth ("Garbage")	Time Aggregated ("NIPA")	Unfiltered NIPA Consumption				
			P-J	Q4Q4	$\phi = .80$	$\phi = .53$	$\phi = .40$
Mean %	1.05	1.03	2.99	1.04	1.05	1.11	1.21
St. dev. %	3.01	2.45	4.79	2.90	2.99	4.64	6.40
Autocorr.	-0.02	0.22	0.72	0.02	0.05	-0.17	-0.24
Corr. R^e_M	0.57	0.37	0.34	0.54	0.39	0.38	0.37
Cov. R^e_M $\times 100$	0.36	0.19	0.35	0.33	0.24	0.37	0.49
RRA γ	15.04	27.62	15.84	16.40	21.95	14.15	10.43

Panel B	Time Aggregation and Filtering ($\nu = 1/2$)						
	Consumption Growth ("Garbage")	Time Aggr. & Filtered ("NIPA")	Unfiltered NIPA Consumption				
			P-J	Q4Q4	$\phi = .80$	$\phi = .53$	$\phi = .40$
Mean %	1.05	1.01	2.99	1.02	1.01	1.03	1.05
St. dev. %	3.01	1.54	3.77	2.21	1.72	2.33	3.03
Autocorr.	-0.02	0.61	0.84	0.00	0.49	0.25	0.14
Corr. R^e_M	0.57	0.28	0.35	0.49	0.32	0.37	0.38
Cov. R^e_M $\times 100$	0.36	0.09	0.27	0.23	0.12	0.18	0.24
RRA γ	15.04	55.02	20.31	24.10	44.80	29.34	21.88

In Panel A, the properties of NIPA consumption are reasonably close to theoretical predictions. Standard deviation, first-order autocorrelation, and - most importantly - covariation with the stock market deviate from true consumption as can be expected from time-aggregated data (see Section 1.2). As a result, a Generalized Method of Moments (GMM) estimate for the coefficient of relative risk aversion based on the moment condition implied by Equation (10) appears in the simulation with a value of 28, which is about twice as large as the true coefficient of risk aversion of 15. The P-J and the Q4Q4 measures produce a large stock market covariation and thus risk aversion estimates which are close to 15. However, notice that Q4Q4 consumption really improves the correlation, whereas the P-J measure raises covariation through a larger standard deviation.

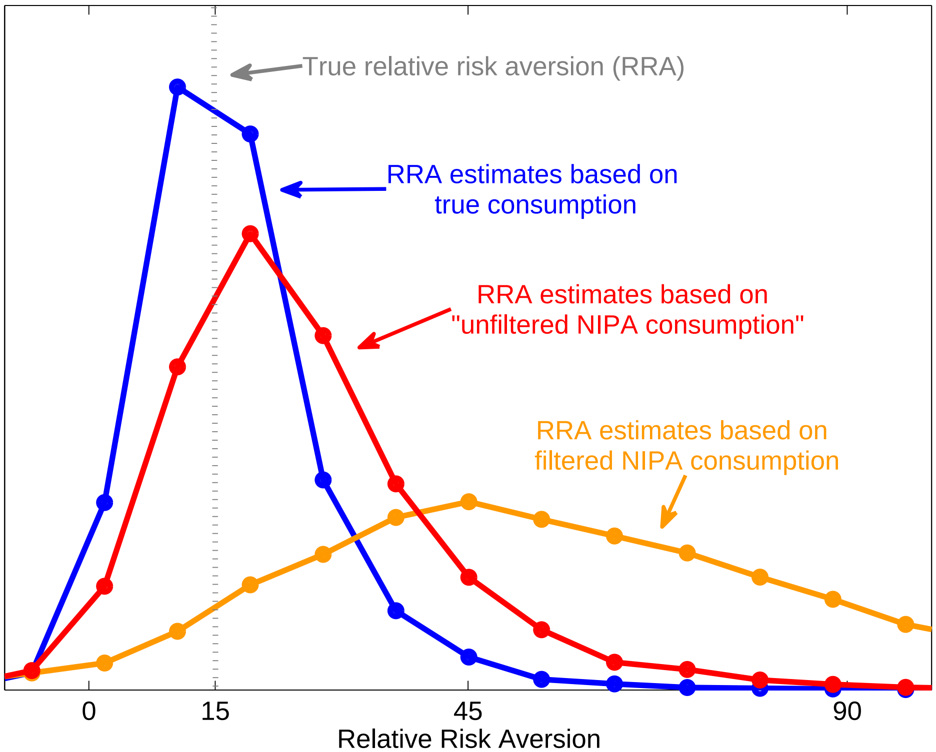
The new method with an unfilter parameter $\phi = .80$ slightly increases stock market correlation compared to NIPA consumption (from 0.37 to 0.39) and recovers the true standard deviation of consumption of 3%. As a result, stock market covariation improves considerably and the GMM estimate of risk aversion is 22, which lies in the middle of the true value and the NIPA consumption-based estimate. Lower values of ϕ raise standard deviation of consumption, but not stock market correlation. As a result, stock market covariation rises, and thus brings down the risk aversion estimate. This finding suggests that the parameter ϕ should be carefully calibrated in empirical work. As a practical guide, the choice of ϕ can be bound by a plausible standard deviation of observable consumption (e.g. garbage) to avoid arbitrary large covariances.

In Panel B, true consumption is time aggregated and filtered ($\nu = 1/2$). The impact on the stochastic properties of consumption are substantial. The standard deviation of simulated NIPA consumption reduces to 1.5%, autocorrelation increases to 0.61, stock market correlation reduces to 0.28, and a tiny covariance of just 0.09 is left. As a result, the GMM estimate of relative risk aversion is 55, which is more than three times larger than the true value of 15.

Again, P-J and Q4Q4 turn out to be reasonable unfilter rules. However, there are important differences to mere time aggregation. The P-J measure increases the stock market correlation of NIPA consumption, whereas the stock market correlation decreases in Panel A. The Q4Q4 measure recovers less of the stock market correlation compared to Panel A, and thus, covariation is relatively low. Overall, Q4Q4 becomes less effective in recovering true consumption when filtering is added.

Figure 1: Relative Risk Aversion Estimates in a Simulated Asset Pricing Economy

The figure shows the simulated distribution of GMM-based relative risk aversion estimates of true consumption, time aggregated and filtered NIPA consumption, and unfiltered ($\phi = .40$) NIPA consumption in a consumption-based asset pricing economy with a true coefficient of relative risk aversion of 15. See Table 1 for details.



Unfiltered NIPA consumption performs fairly. Using the “correct” unfilter parameter value of $\phi = .40$, the standard deviation of unfiltered NIPA consumption is very close to the value of true consumption, and the covariance is 0.24 in contrast to 0.09 for NIPA consumption. The GMM estimate of risk aversion is 22, which is still biased upwards, however, which is now much closer now to true risk aversion compared to estimates based on the original NIPA consumption measure. Figure 1 illustrates this point showing the simulated distribution of relative risk aversion estimates of true consumption, time aggregated and filtered NIPA consumption, and unfiltered NIPA consumption.

The simulation is summarized as follows. Filtering intensified by time aggregation results in substantial distortions of the measured stochastic properties of consumption. The degree of distortion is largely underestimated with respect to the well-known effect of time aggregation. The P-J and the Q4Q4 consumption measures recover some aspects of the stochastic properties of observed consumption, and can thus be viewed as ad hoc unfilter rules. Interestingly, with respect to the stock market covariance of consumption, the Q4Q4 measure is effective in removing the bias stemming from time aggregation, the P-J measure is more effective in removing filtering. However, both ad hoc unfilter rules have difficulties in matching other aspects of the data. The new method to unfilter NIPA consumption performs well across all stochastic properties and GMM estimates of risk aversion are conservative as long as the unfilter parameter ϕ is reasonably calibrated.

1.5 Asset Pricing without Garbage

Filtering can be harmful for consumption-based asset pricing, and may explain the relative success of alternative consumption measures. Thus, I calibrate the structural model of Section 1.2 to bring unfiltered NIPA consumption close to garbage. Afterwards, I will compare unfiltered NIPA consumption, garbage, P-J, and Q4Q4 in popular asset pricing applications.

1.5.1 Comparison of Alternative Consumption Measures

Panel A of Table 2 shows well-known statistics for NIPA annual real per capita consumption growth based on nondurables and services, as well as for nondurables and services separately. Details on data sources and construction is provided in Appendix A.2.[14] Results are reported for the full sample period from 1927 to 2011 and the post-war sample from 1950 to 2011. I treat garbage as observable consumption, which should not be subject to filtering or time aggregation.[15] Data on garbage is only available for the post-war period, for the shorter sampler

[14]Throughout the study, I use the "end-of-period" timing convention for all NIPA consumption measures, as explained in Appendix A.2.

[15]The assumption that garbage is not filtered is in line with the empirical properties of garbage and is based on the estimation methodology of garbage discussed in the internet appendix to Savov (2011). The assumption that garbage is not subject to time aggregation is more difficult to make. However, the empirical properties of garbage hardly point to a time aggregation effect (this observation is also mentioned in the internet appendix to Savov (2011)). Treating garbage as not time aggregated for the calibration of the unfilter rule is conservative in

period 1960 to 2011. I use the empirical moments of garbage to calibrate unfiltered NIPA consumption using the parameter ϕ as in Equation (9).

any case. Otherwise the true covariance of stock returns with consumption would be larger by a factor of two and the correlation by a factor of one and a half.

Table 2: Empirical Properties of Alternative Consumption Measures

The table reports empirical moments of U.S. annual real per capita consumption measures. Nondurables and services are obtained from NIPA. Statistics on Garbage are are based on a slightly shorter sample period from 1960 to 2007 . Panel B and Panel C show three unfilter rules for the NIPA consumption measure. P-J is three-year NIPA consumption growth (Parker and Julliard (2005)). Q4Q4 is fourth-quarter to fourth-quarter NIPA consumption growth (Jagannathan and Wang (2007)). The third is the direct unfilter rule, as in Equation (9), using two different values for the unfilter parameter ϕ. The lower value of ϕ is chosen to calibrate unfiltered NIPA consumption to match garbage in the post-war sample. The other value of ϕ serves as a sensitivity check.

Panel A	NIPA Consumption and Garbage						
	Full Sample (1927 - 2011)			Post-War Sample (1950 - 2011)			
	NIPA Consumption				NIPA Consumption		
	Nond. & Serv.	Nond.	Serv.	Garbage	Nond. & Serv.	Nond.	Serv.
Mean %	1.81	1.38	2.13	1.47	1.91	1.34	2.31
St. dev. %	2.22	2.65	2.15	2.88	1.30	1.55	1.34
Autocorr.	0.39	0.27	0.52	-0.14	0.43	0.25	0.54
Corr. R^e_M	0.12	0.17	0.08	0.58	0.01	0.09	-0.01
Cov. R^e_M ×100	0.06	0.10	0.03	0.27	0.00	0.03	-0.00

Panel B	Unfiltered NIPA Consumption: Nondurables & Services						
	Full Sample (1927 - 2011)			Post-War Sample (1950 - 2011)			
	P-J	$\phi = .53$	$\phi = .40$	P-J	Q4-Q4	$\phi = .53$	$\phi = .40$
Mean %	5.54	1.75	1.71	5.75	1.91	1.92	1.93
St. dev. %	4.99	3.93	5.35	3.07	1.48	2.22	3.00
Autocorr.	0.77	-0.05	-0.15	0.80	0.31	0.01	-0.09
Corr. R^e_M	0.41	0.20	0.21	0.18	0.34	0.17	0.21
Cov. R^e_M ×100	0.42	0.16	0.24	0.10	0.09	0.07	0.11

Panel C	Unfiltered NIPA Consumption: Nondurables excluding Services						
	Full Sample (1927 - 2011)			Post-War Sample (1950 - 2011)			
	P-J	$\phi = .80$	$\phi = .53$	P-J	Q4-Q4	$\phi = .80$	$\phi = .53$
Mean %	4.20	1.36	1.32	4.02	1.33	1.34	1.36
St. dev. %	5.54	3.21	4.96	3.13	1.95	1.88	2.90
Autocorr.	0.71	0.08	-0.15	0.74	0.12	0.09	-0.13
Corr. R^e_M	0.43	0.22	0.26	0.23	0.41	0.15	0.23
Cov. R^e_M ×100	0.49	0.14	0.26	0.13	0.14	0.05	0.12

In Panel A, focusing on the post-war sample at the right hand side of the table, the standard deviation of reported NIPA nondurables and services is less than one-half the standard deviation of garbage, and the first-order autocorrelation is about 0.57 larger. The stock market covariance of the traditional NIPA measure is virtually zero, in contrast to 0.27 for garbage. These two findings combined can hardly be explained by measurement error or time aggregation, but are reasonably in line with an additional effect of filtering, as documented in the simulation experiment of the previous section. There is also evidence of filtering when looking separately on NIPA nondurables and NIPA services. The standard deviation of NIPA services is lower than the standard deviation of NIPA nondurables, autocorrelation is larger, and the measured stock market covariance is smaller (even negative). The time aggregation bias should be the same for both time series. However, these differences are plausible if consumption data is filtered. If measurement error for the individual components of NIPA services is larger than for NIPA nondurables, the parameter ϕ should be lower for NIPA services, implying a more aggressive filter. There is ample evidence in the official BEA documentations that NIPA services are indeed more imprecisely measured, and NIPA services are more often subject to interpolation (BEA (2009)).

Panel B compares several unfilter rules for NIPA nondurables and services, the most common consumption measure in asset pricing. Using Q4Q4 and P-J as unfilter rules increases the stock market covariance of consumption from zero to 0.09 and 0.10 in the post-war sample. Stock market correlation of the Q4Q4 measure is larger than for the P-J measure. However, the standard deviation of the Q4Q4 measure is lower, netting to similar covariances. Again, this pattern is in line with the simulation experiment of the previous section, when reported consumption is subject to filtering. Unfiltered NIPA consumption is able to match garbage quite closely with a calibrated value of $\phi = .40$. This corresponds to a pure filter ν of 1/2, after accounting for time aggregation ($\phi = .80 \times \nu$). The covariance of unfiltered NIPA nondurables and services is about 0.11 and larger than for the P-J measure and the Q4Q4 measure.

Wilcox (1992) concludes that “from the perspective of the measurement system, the usual two-part disaggregation (motivated from theory) into (i) durables and (ii) nondurables plus services does not make much sense”. Wilcox suggests to distinguish between nondurables and services in empirical work. Similarly, Savov (2011) suspects that particularly the way in which services are measured and estimated is rather harmful for the NIPA measure and may (partly) explain differences to garbage. For this reason, I also consider NIPA nondurables excluding services in Panel C.[16] In comparison to Panel B, I find that standard deviations increase, autocorrelations decrease, and most importantly, correlations and covariances to the stock market further increase for the P-J and Q4Q4 measures. Interestingly, the qualitative improvements are once again in line with the Monte Carlo experiment. As an example, when filtering is likely to be lower, as in Panel C, the improvement to the covariance of the Q4Q4 measure is large relatively to the P-J measure. The direct unfilter method is able to match the properties of garbage closely for a more moderate unfilter parameter ϕ of only .53 (corresponding to $\nu = 2/3$, as in the simulation). Again, a higher value of ϕ is in line with less measurement error in NIPA nondurables than in NIPA services.

Finally, it is noted that in the post-war sample, stock market covariances of reported NIPA consumption (Panel A) as well as the alternative NIPA-based consumption measures (Panel B and Panel C) are lower than one might expect from the simulation experiment. However, this is less the case in the longer sample from 1927 to 2011, where covariances are several times larger for all NIPA-based consumption measures. Time series plots of the alternative consumption measures are provided in Figure 2.

In summary, I find that time aggregation does not suffice to explain the empirical properties of NIPA consumption as well as alternative measures such as P-J and Q4Q4. Although the filter proposed in this paper is a clear simplification of reality, it is sufficient to match important aspects of the data surprisingly well and brings the properties of unfiltered NIPA consumption close to garbage.

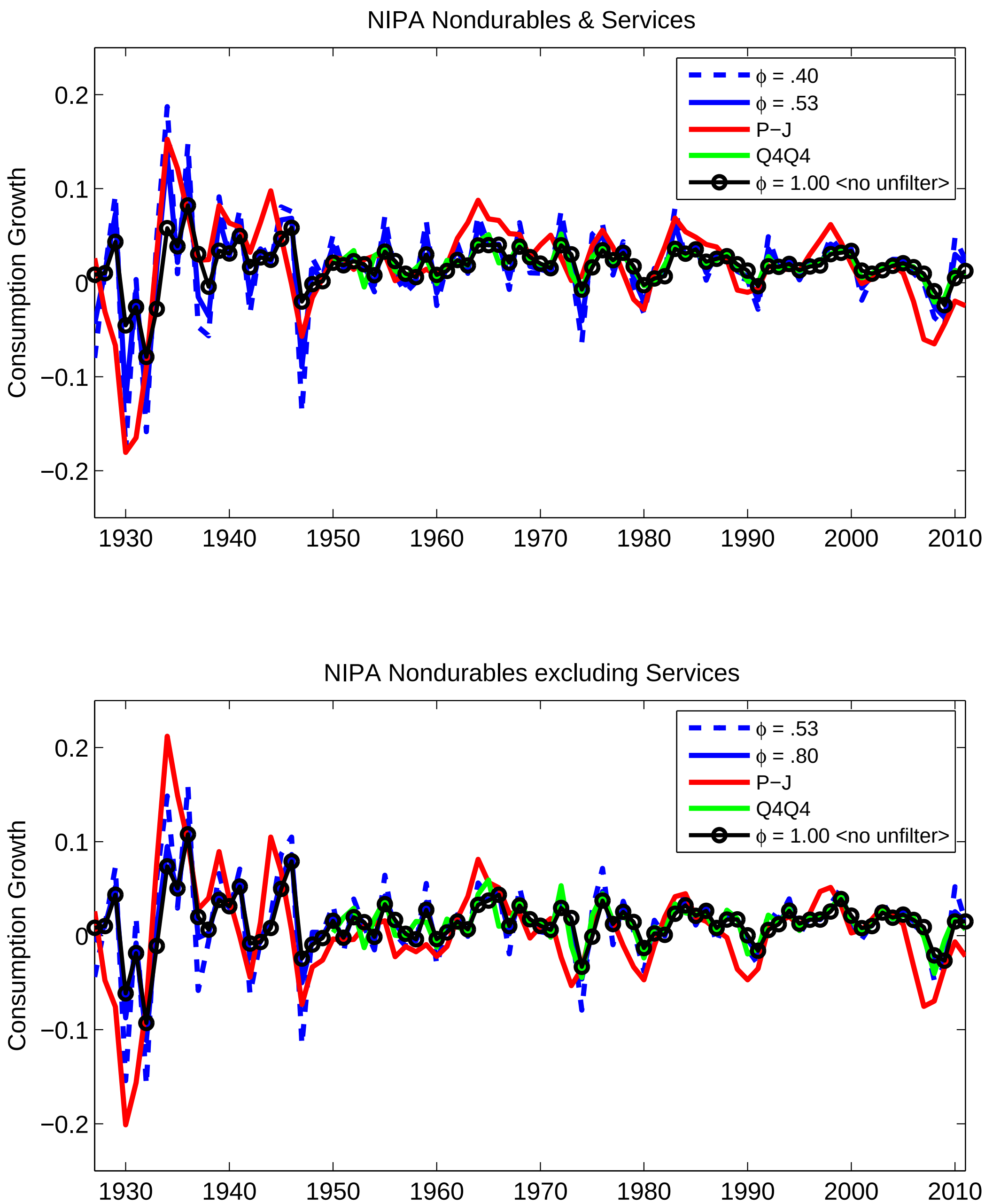
1.5.2 Equity Premium

Model. I test the Euler equation implied by the consumption-based asset pricing model (Lucas (1978), Breeden (1979)) as stated in Equation (10). I follow Savov (2011), among others, and fix $\beta = .95$ to focus on relative risk aversion. It is well known that estimates of the relative risk

[16] Hall (1988), among others, tests nondurables excluding services as well.

Figure 2: NIPA Consumption Measures

The figure shows annual real per capita NIPA consumption growth from 1927 to 2011. The upper figure shows NIPA nondurables and services, and the lower figure shows NIPA nondurables excluding services. The consumption measures are NIPA consumption without unfaltering, unfiltered NIPA consumption, three-year NIPA consumption (P-J) as in Parker and Julliard (2005), and fourth-quarter to fourth-quarter NIPA consumption (Q4Q4) as in Jagannathan and Wang (2007).



aversion obtained from this Euler equation are typically too high from a theoretical perspective (Hansen and Singleton (1982), Mehra and Prescott (1985)). The risk-free rate puzzle, termed by Weil (1989), is related to the equity premium puzzle. High coefficients of relative risk aversion imply a high desire for consumption smoothing, resulting in implausibly large unconditional real risk-free interest rates. A log approximation of Equation (10) yields for the risk-free rate (e.g., Campbell, Lo, and MacKinlay (1997)):

$$r_f = -log(\beta) + \gamma E\left[log(C_{t+1}/C_t)\right] - \frac{1}{2}\gamma^2 Var\left[log(C_{t+1}/C_t)\right]. \tag{11}$$

Following Savov (2011) when testing Equation (10), I also report this implied risk-free rate. Furthermore, the empirically observed risk-free rate can be directly included as a test asset in addition to Equation (10):

$$E\left[\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} R_{f,t+1} - 1\right] = 0, \tag{12}$$

where the estimation procedure is forced to match the equity premium jointly with the empirical risk-free rate. Estimation is carried out by Generalized Method of Moments (GMM) with Newey and West (1987) standard errors (three lags, and removed moment means). I report the mean absolute error (MAE) of the moment condition(s), to check if the Euler equation errors are economically large (Lettau and Ludvigson (2009)), and the p-value to the J-test of overidentified restrictions.

Results: NIPA Nondurables and Services. Table 3 provides GMM estimates for the coefficient of relative risk aversion for unfiltered NIPA nondurables and services. The table is split into a full sample period from 1927 to 2011 and a shorter post-war sample period from 1950 to 2011. Garbage and the Q4Q4 measure are only available in the post-war sample. As shown in the previous section, the calibrated unfilter parameter which brings NIPA nondurables and services close to garbage is equal to $\phi = .40$ ($\nu = 1/2$). I also report results for a larger parameter, $\phi = .53$ ($\nu = 2/3$), to provide some comparative analysis on the effect of the unfilter parameter. Finally, the table also provides results for the traditional NIPA consumption measure ($\phi = 1$).

In Panel A, the equity premium is the only test asset. In the full sample, the traditional NIPA consumption measure requires a coefficient of relative risk aversion of 34, with an implied risk-free rate of 38%. In contrast, unfiltered NIPA consumption ($\phi = .40$) requires a low estimate of 12 with an implied risk-free rate of only 5%. The ultimate consumption risk measure proposed by Parker and Julliard (2005) produces a relatively low coefficient of relative risk aversion of 9, but goes hand in hand with a large risk-free rate of 45%.

Related research (e.g. Campbell (2003), Engsted and Møller (2011)) has shown that excluding the turbulent periods in the first half of the 20th century substantially increases the equity premium (and risk-free rate) puzzle. In line with the literature, as shown in Table 3, the traditional NIPA consumption measure requires a relative risk aversion of 59 with a risk-free rate of 89% in the post-war period and still produces a MAE of 3.2%. In contrast, garbage as an alternative consumption measure matches the equity premium with a relative risk aversion of only 15 and an implied risk-free rate of 18%. The P-J and Q4Q4 measures perform better than traditional NIPA consumption, but still substantially worse than garbage. The table shows that both measures have difficulties particularly with matching the risk-free rate. I find that unfiltered NIPA consumption is closer to garbage. When setting the parameter ϕ to .40, the coefficient of relative risk aversion in the post-war sample is 34, which is within one standard error to 15. Similarly, the implied risk-free rate of unfiltered NIPA consumption is 19%, close to garbage as well (18%). In the post-war period, the estimated coefficients of relative risk aversion for unfiltered NIPA consumption and garbage are still high but much smaller than for the other NIPA-based consumption measures.

Panel B forces GMM to fit both, the equity premium and the empirical risk-free rate at the same time. The traditional NIPA consumption measure, the P-J measure, and the Q4Q4 measure fail to fit both moment conditions even with a large relative risk aversion, as indicated by a large MAE. In contrast, unfiltered NIPA consumption performs considerably better. Relative

Table 3: GMM Estimates of Relative Risk Aversion

The table reports GMM estimates of the relative risk aversion coefficient γ (RRA). The moment restrictions $g(\gamma)$ are:

$$E\left[\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} R^e_{m,t+1} - 0\right] \text{ and } E\left[\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} R_{f,t+1} - 1\right],$$

where R^e_m is the market excess return and R_f is the gross risk-free rate. The discount parameter is fixed, $\beta = 0.95$. Panel A uses only the first moment restriction (the equity premium), where r_f is the implied log risk-free rate. Panel B exploits both moment restrictions. The consumption measures are reported NIPA consumption ($\phi = 1$), unfiltered NIPA consumption ($\phi = .53$, and $\phi = .40$), three-year NIPA consumption (P-J) as in Parker and Julliard (2005), and fourth-quarter to fourth-quarter NIPA consumption (Q4Q4) as in Jagannathan and Wang (2007). Estimates on Garbage are are based on a sample period from 1960 to 2007. MAE is the mean absolute error of the moment restrictions. Panel A is exactly identified. The p-values in Panel B correspond to the J-test of overidentified restrictions.

	Full Sample (1927 - 2011)				Post-War Sample (1950 - 2011)					
	NIPA Nondurables & Services				Garbage	NIPA Nondurables & Services				
	$\phi = 1$	P-J	$\phi = .53$	$\phi = .40$		$\phi = 1$	P-J	Q4Q4	$\phi = .53$	$\phi = .40$
Panel A	Equity Premium									
RRA (γ)	34.37	9.07	16.82	12.14	14.92	59.32	35.26	55.78	56.67	34.34
(s.e.)	(14.12)	(4.47)	(6.99)	(5.12)	(8.35)	-	(16.57)	(27.24)	(48.88)	(20.68)
Implied r_f, %	38.13	45.16	12.76	4.84	17.80	88.59	149.26	77.37	34.84	18.51
MAE %	0.00	0.00	0.00	0.00	0.00	3.20	0.00	0.00	0.00	0.00
Panel B	Equity Premium & Risk-free Rate									
	One-Stage GMM									
RRA (γ)	41.13	20.89	16.06	9.28	25.37	157.19	126.21	163.79	66.16	38.04
(s.e.)	(13.71)	(7.03)	(6.07)	(3.86)	(11.02)	(42.07)	(34.21)	(44.98)	(16.90)	(11.04)
MAE %	1.61	14.08	0.34	1.59	3.51	11.80	12.85	8.62	0.83	0.69
p-value (J)	0.80	0.05	0.96	0.81	0.71	0.13	0.10	0.25	0.90	0.92
	Two-Stage GMM									
RRA (γ)	21.03	2.43	16.09	8.86	7.57	130.92	104.68	151.46	66.37	38.47
(s.e.)	(10.22)	(0.31)	(6.07)	(3.94)	(2.67)	(42.17)	(32.98)	(44.75)	(16.82)	(10.78)
MAE %	12.36	8.47	0.36	1.50	4.72	23.59	27.29	14.50	0.92	1.01
p-value (J)	0.59	0.79	0.83	0.26	0.66	0.00	0.00	0.09	0.86	0.83

risk aversion coefficients are lower, and the MAE of the moment conditions is relatively small. In particular, results for the full sample show that unfiltered NIPA consumption requires a relative risk aversion as small as 9. In the post-war sample, relative risk aversion is higher, around 38, which is nevertheless four times lower than for traditional NIPA consumption, and thus again much closer to garbage.

Results: NIPA Nondurables Excluding Services. Given the measurement methodology of consumption, Wilcox (1992) argues that the services component in NIPA consumption could be harmful for empirical applications. To quantify the impact of services, Table 4 provides results for alternative consumption measures based on NIPA nondurables excluding services. Less measurement error also means less filtering, and thus, allows a larger value for the unfilter parameter. As shown in the previous section, the calibrated unfilter parameter which brings NIPA nondurables excluding services close to garbage is $\phi = .53$ ($\nu = 2/3$). For comparison, the table also provides results for $\phi = .80$ ($\nu = 1$).

Overall, I find that excluding services significantly improves the results of all NIPA-based alternative consumption measures, particularly in the post-war sample (1950 - 2011). For example, the Q4Q4 measure results in a coefficient of relative risk aversion of only 30, in contrast to 56 when services are included (Panel A). However, the P-J as well as the Q4Q4 measures both still experience problems fitting the risk-free rate. In contrast, using a modest parameter $\phi = .53$, unfiltered NIPA nondurables require a coefficient of relative risk aversion of only 26 with an implied risk-free rate of 12%, which is indistinguishable from garbage. In the long sample (1927 - 1951), I find a coefficient of relative risk aversion of 13 for unfiltered NIPA consumption, with an implied (real) risk-free rate of 2%.

To summarize, I find that after accounting for filtering, it is possible to replicate the results of garbage relying only on NIPA consumption. For the longer sample starting in 1927, I find an even lower coefficient of relative risk aversion between 9 and 13 with a low implied risk-free rate between 2% and 5% which is close to theoretically plausible values.

Table 4: GMM Estimates of Relative Risk Aversion: Excluding Services

The table reports GMM estimates of the relative risk aversion coefficient γ (RRA). The moment restrictions $g(\gamma)$ are:

$$E\left[\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} R^e_{m,t+1} - 0\right] \text{ and } E\left[\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} R_{f,t+1} - 1\right],$$

where R^e_m is the market excess return and R_f is the gross risk-free rate. The discount parameter is fixed, $\beta = 0.95$. Panel A uses only the first moment restriction (the equity premium), where r_f is the implied log risk-free rate. Panel B exploits both moment restrictions. The consumption measures are reported NIPA consumption ($\phi = 1$), unfiltered NIPA consumption ($\phi = .80$, and $\phi = .53$), three-year NIPA consumption (P-J) as in Parker and Julliard (2005), and fourth-quarter to fourth-quarter NIPA consumption (Q4Q4) as in Jagannathan and Wang (2007). Estimates on Garbage are based on a sample period from 1960 to 2007. MAE is the mean absolute error of the moment restrictions. Panel A is exactly identified. The p-values in Panel B correspond to the J-test of overidentified restrictions.

	Full Sample (1927 - 2011)				Post-War Sample (1950 - 2011)					
	NIPA Nondurables excluding Services				Garbage	NIPA Nondurables excluding Services				
	$\phi = 1$	P-J	$\phi = .80$	$\phi = .53$		$\phi = 1$	P-J	Q4Q4	$\phi = .80$	$\phi = .53$
Panel A	Equity Premium									
RRA (γ)	27.84	9.10	21.17	12.80	14.92	69.18	32.77	30.36	46.28	26.47
(s.e.)	(10.75)	(4.55)	(8.24)	(5.08)	(8.35)	(41.38)	(15.53)	(13.25)	(23.08)	(12.69)
Implied r_f, %	16.17	30.63	10.89	1.92	17.80	40.12	84.38	28.06	29.54	11.60
MAE %	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Panel B	Equity Premium & Risk-free Rate									
	One-Stage GMM									
RRA (γ)	28.72	17.30	20.23	9.18	25.37	83.97	85.70	53.10	57.56	26.43
(s.e.)	(10.87)	(6.53)	(7.96)	(3.92)	(11.02)	(22.90)	(23.90)	(16.36)	(17.38)	(9.46)
MAE %	0.24	8.70	0.32	1.79	3.51	1.60	5.63	5.53	2.09	0.01
p-value (J)	0.97	0.23	0.96	0.79	0.71	0.83	0.43	0.50	0.79	1.00
	Two-Stage GMM									
RRA (γ)	28.16	3.22	20.53	8.82	7.57	84.88	8.34	11.30	58.91	26.42
(s.e.)	(10.39)	(0.60)	(7.74)	(3.99)	(2.67)	(22.62)	(1.38)	(2.67)	(17.03)	(9.37)
MAE %	0.65	7.84	0.54	1.56	4.72	1.95	14.61	7.76	2.62	0.01
p-value (J)	0.89	0.74	0.83	0.16	0.66	0.71	0.68	0.61	0.49	1.00

1.5.3 Cross-Section of Stock Returns

Model. A linearized version of the consumption-based asset pricing model implies the following (cross-sectional) beta representation (Breeden, Gibbons, and Litzenberger (1989)):

$$E\left[R^e_{i,t+1}\right] = \lambda_0 + \lambda\beta_i, \tag{13}$$

where $E\left[R^e_{i,t+1}\right]$ is the expected excess return of asset i, and λ_0 is a common pricing error which should be zero by theory. The 25 Fama-French portfolios sorted by size and book-to-market, available on the web site of Kenneth R. French, are used as test assets. The two-stage cross-sectional regression method of Fama and MacBeth (1973) can be used to estimate the specification in Equation (13). In a first step, consumption growth betas are estimated using a time-series OLS regression for each asset i. In a second step, a cross-sectional OLS regression on the first step betas is used to estimate the risk factor price λ.

The cross-sectional OLS estimates do not account for the fact that the betas themselves are estimated, nor do they account for heteroskedasticity and autocorrelation. Shanken (1992) provides adjustments for the standard errors with respect to the errors in variables problem. Cochrane (2005) shows how to use the GMM procedure to account for both issues. For GMM-based inference, I apply the parametric VARHAC method described by den Haan and Levin (2000). Burnside (2011) argues that if some factors are persistent, as is particularly the case for the P-J measure, the GMM-VARHAC should be preferred to the Newey-West approach to calculate standard errors.[17] The weighting matrix is set up such that the point estimates of the GMM procedures are identical to the traditional Fama-MacBeth regressions (details are provided in Cochrane (2005), and Burnside (2011)).

I report the cross-sectional R^2, the mean absolute error (MAE), and a χ^2-test on the 25 pricing errors of the model (Cochrane (2005, p.234)) as measures of model fit.

[17] To implement the GMM-VARHAC, I apply the VARHAC matlab code with the same settings as Burnside (2011) available to the online appendix of Burnside's paper on http://www.aeaweb.org.

Results: NIPA Nondurables and Services. Table 5 shows the cross-sectional regression results for NIPA nondurables and services using the full sample (1927 - 2011) and the shorter post-war sample (1950 - 2011). The discussion focuses on the post-war sample. Simple NIPA consumption ($\phi = 1$) as a risk factor results in a large and significant constant (9.5%), and a low cross-sectional fit ($R^2 = 0.28$). Garbage has difficulties in pricing the 25 Fama-French portfolios, and shows no advantage compared to NIPA consumption when a constant is included to the model. The P-J and Q4Q4 measures perform better. The constants (λ_0) are insignificant based on GMM-VARHAC t-statistics, close to zero for the Q4Q4 measure, and risk factors (λ) are significant for the alternative NIPA-based consumption measures. For both, the one-factor consumption model explains more than one-half of the cross-sectional variation of average returns.

The new consumption measure, unfiltered NIPA consumption, performs similarly to P-J and Q4Q4. Using the calibrated unfilter parameter $\phi = .40$, the constant is not significant and the risk factor is significantly priced in the cross-section of stock returns with a slope coefficient of 4.3%. The R^2 is about 0.59, which is as large as for the P-J and Q4Q4 measures. The mean absolute error of unfiltered NIPA consumption is 1.4%, close to P-J (1.2%) and Q4Q4 (1.3%). However, based on GMM inference, all models are rejected by the χ^2 test on the 25 pricing errors at the 5% level. At the bottom of the table, the cross-sectional regressions are repeated with the restriction of a zero constant. I find that in the post-war sample, consumption is only significantly (5% level) priced when using the Q4Q4 measure, and, in contrast to when a constant is included, for the garbage measure. In the longer sample period, the estimated risk premium is significant for all considered consumption measures with the exception of traditional NIPA consumption ($\phi = 1$).

Results: NIPA Nondurables Excluding Services. Table 6 provides the cross-sectional regression results for NIPA nondurables excluding services. I find that excluding services improves some aspects of the cross-sectional model, particularly in the post-war sample from 1950 to 2011. Common pricing errors and MAEs are lower and the cross-sectional R^2 is larger for almost all

Table 5: Fama-MacBeth Estimates: NIPA Nondurables & Services

The table displays estimates on the consumption risk premium (λ) from the cross-sectional regression of 25 Fama-French portfolios:

$$\bar{R}^e_{i,t} = \lambda_0 + \lambda\beta_i + u_i$$

where $\bar{R}^e_{i,t}$ is mean excess return of portfolio i, λ_0 is a constant (common pricing error), and β_i is the time-series beta of a first pass regression of the portfolio return on a consumption measure based on NIPA nondurables (excluding services). The consumption measures are NIPA consumption ($\phi = 1$), unfiltered NIPA consumption ($\phi = .53$, and $\phi = .40$), three-year NIPA consumption (P-J) as in Parker and Julliard (2005), and fourth-quarter to fourth-quarter NIPA consumption (Q4Q4) as in Jagannathan and Wang (2007).

	Full Sample (1927-2011)				Post-War Sample (1950 - 2011)					
	NIPA Nondurables & Services					NIPA Nondurables & Services				
	$\phi = 1$	P-J	$\phi = .53$	$\phi = .40$	Garbage	$\phi = 1$	P-J	Q4Q4	$\phi = .53$	$\phi = .40$
λ_0	10.16	1.99	1.56	0.25	8.80	9.47	4.65	-0.07	5.05	3.74
$t_{shanken}$	(3.54)	(0.71)	(0.37)	(0.06)	(2.10)	(2.37)	(1.23)	(-0.01)	(0.94)	(0.65)
t_{varhac}	[3.62]	[0.64]	[0.34]	[0.06]	[2.09]	[1.84]	[0.92]	[-0.01]	[0.70]	[0.49]
λ	-0.08	3.40	5.83	8.31	-0.12	1.47	3.47	1.94	3.06	4.27
$t_{shanken}$	(-0.15)	(2.20)	(2.06)	(2.12)	(-0.08)	(1.46)	(2.33)	(2.14)	(1.77)	(1.84)
t_{varhac}	[-0.14]	[2.68]	[2.63]	[3.23]	[-0.08]	[1.50]	[2.51]	[2.18]	[1.75]	[1.81]
R^2	0.00	0.30	0.43	0.57	0.00	0.28	0.56	0.58	0.54	0.59
MAE	2.63	1.93	1.83	1.61	2.73	1.84	1.23	1.32	1.49	1.40
χ^2	(0.000)	(0.000)	(0.341)	(0.420)	(0.000)	(0.022)	(0.033)	(0.130)	(0.154)	(0.202)
	[0.000]	[0.000]	[0.005]	[0.002]	[0.000]	[0.000]	[0.027]	[0.006]	[0.012]	[0.007]
	Estimation without Intercept ($\lambda_0 = 0$)									
λ	6.35	4.20	6.85	8.51	2.27	1.12	6.35	1.93	6.00	6.75
$t_{shanken}$	(1.12)	(2.74)	(1.77)	(1.90)	(2.38)	(1.22)	(1.65)	(2.31)	(1.37)	(1.58)
t_{varhac}	[0.77]	[2.78]	[1.79]	[2.06]	[2.30]	[1.17]	[1.61]	[1.96]	[1.14]	[1.32]

NIPA-based consumption measures. For unfiltered NIPA consumption ($\phi = .53$), the constant is only 0.9% and the R^2 is 0.62. However, based on GMM inference and when estimation is without an intercept and the sample period is restricted to 1950-2011, the risk factor premium is not significant (5% level) for unfiltered NIPA nondurables. Again, taking the full sample period from 1927 to 2011 leads to t-statistics well above two and points to significance.

Consumption Betas. It is useful to verify the cross-sectional regression results on the consumption risk premium with a look at the first-step consumption betas, or risk exposures, of the individual test assets. Small stocks have larger average returns than big stocks. High book-to-market stocks (value stocks) have larger average returns than low book-to-market stocks (growth stocks). Accordingly, the first step consumption betas should have respective spreads to provide a reasonable explanation for differences in average returns.

Figure 3 provides the underlying consumption betas of Table 5 and Table 6, for reported NIPA consumption ($\phi = 1$) and unfiltered NIPA consumption ($\phi = \{.53, .40\}$ for NIPA nondurables and services and $\phi = \{.80, .53\}$ for NIPA nondurables excluding services). Overall, the consumption betas for reported NIPA consumption hardly show any systematic size or value exposure. However, the unfiltered NIPA consumption measures show a smooth ramp up from big to small stocks and from growth to value stocks. Filtering and time aggregation in consumption data seem to hide the consumption risk exposure of the Fama-French portfolio returns. The simple unfilter formula of Equation (9) goes a long way in recovering betas which are in line with theory.

1.5.4 The Best of two Measures: A Combined Approach

It is probably possible to further improve the simple unfilter approach proposed in this paper. As a final remark, I consider a straightforward modification in this section. The simulation experiment provides evidence that using fourth-quarter to fourth-quarter consumption is effective in dissolving effects of time aggregation. In contrast, the unfilter method is more effective in dissolving the effects of filtering. Thus, a combination of both "unfilter rules" is promising.

Table 6: Fama-MacBeth Estimates: NIPA Nondurables Excluding Services

The table reports estimates on the consumption risk premium (λ) from the cross-sectional regression of 25 Fama-French portfolios:

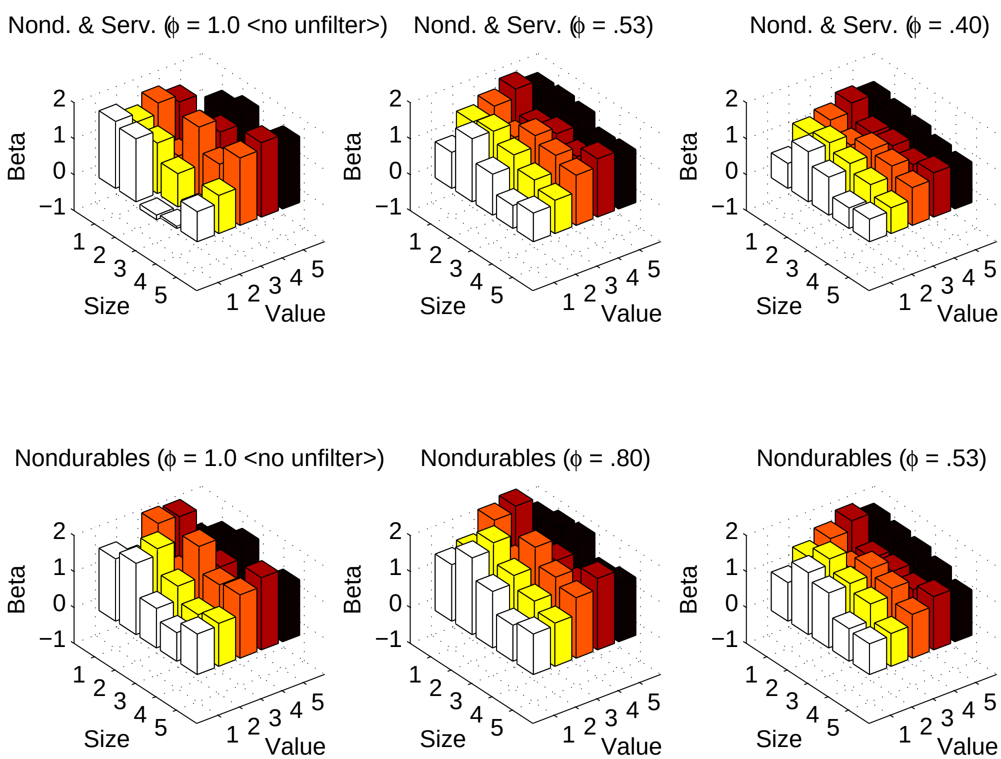
$$\bar{R}^e_{i,t} = \lambda_0 + \lambda\beta_i + u_i$$

where $\bar{R}^e_{i,t}$ is mean excess return of portfolio i, λ_0 is a constant (common pricing error), and β_i is the time-series beta of a first pass regression of the portfolio return on a consumption measure. The consumption measures are NIPA consumption ($\phi = 1$), unfiltered NIPA consumption ($\phi = .80$, and $\phi = .53$), three-year NIPA consumption (P-J) as in Parker and Julliard (2005), and fourth-quarter to fourth-quarter NIPA consumption (Q4Q4) as in Jagannathan and Wang (2007).

	Full Sample (1927-2011)				Post-War Sample (1950 - 2011)					
	NIPA Nondurables excluding Services					NIPA Nondurables excluding Services				
	$\phi = 1$	P-J	$\phi = .80$	$\phi = .53$	Garbage	$\phi = 1$	P-J	Q4Q4	$\phi = .80$	$\phi = .53$
λ_0	8.60	1.45	3.54	-0.14	8.80	6.62	2.90	-2.33	3.62	0.94
$t_{shanken}$	(3.06)	(0.52)	(1.01)	(-0.03)	(2.10)	(1.00)	(0.79)	(-0.44)	(0.49)	(0.13)
t_{varhac}	[2.64]	[0.46]	[0.87]	[-0.03]	[2.09]	[0.80]	[0.59]	[-0.32]	[0.39]	[0.09]
λ	0.94	3.66	3.70	6.88	-0.12	3.13	3.50	2.67	4.05	5.75
$t_{shanken}$	(1.27)	(2.40)	(2.04)	(2.16)	(-0.08)	(1.35)	(2.33)	(1.96)	(1.49)	(1.74)
t_{varhac}	[1.53]	[2.72]	[2.22]	[3.02]	[-0.08]	[1.35]	[2.70]	[2.10]	[1.52]	[1.63]
R^2	0.02	0.36	0.21	0.49	0.00	0.36	0.67	0.59	0.53	0.62
MAE	2.53	1.77	2.17	1.77	2.73	1.84	0.99	1.31	1.49	1.29
χ^2	(0.000)	(0.000)	(0.058)	(0.229)	(0.000)	(0.811)	(0.038)	(0.220)	(0.888)	(0.788)
	[0.000]	[0.000]	[0.002]	[0.000]	[0.000]	[0.073]	[0.000]	[0.000]	[0.030]	[0.027]
	Estimation without Intercept ($\lambda_0 = 0$)									
λ	6.06	4.25	5.62	6.79	2.27	8.68	4.90	2.15	6.32	6.36
$t_{shanken}$	(1.39)	(2.86)	(1.74)	(2.11)	(2.38)	(0.69)	(2.04)	(2.54)	(1.08)	(1.57)
t_{varhac}	[1.21]	[2.80]	[1.77]	[2.33]	[2.30]	[0.62]	[1.91]	[2.29]	[0.91]	[1.24]

Figure 3: Consumption Betas

The figure shows time-series betas of a first pass regression of 25 Fama-French size and book-to-market portfolio returns on a consumption measure based on NIPA nondurables and services and NIPA nondurables (excluding services). The consumption measures are reported NIPA consumption ($\phi = 1.0$) and unfiltered NIPA consumption; with $\phi = .53$, $\phi = .40$ for NIPA nondurables and services, and $\phi = .80$, $\phi = .53$ for NIPA nondurables (excluding services). The estimation period is from 1927 to 2010.



I unfilter fourth-quarter NIPA consumption using Equation (9). Since the use of fourth-quarter to fourth-quarter NIPA consumption already accounts for time aggregation, I only need to account for filtering. For this purpose, I use ν as implied by the calibration of Table 2 but set $a = 1$, which gives $\phi = 1 \times 1/2$ for NIPA nondurables and services and $\phi = 1 \times 2/3$ for NIPA nondurables excluding services. The resulting standard deviation of unfiltered fourth-quarter NIPA nondurables and services (and NIPA nondurables excluding services) is 2.9% (3.0%), with an autocorrelation of -0.07 (-0.10), and a stock market covariance of 0.27 (0.27). These stochastic properties are very close to garbage (see Table 2).

Table 7 shows that it is possible to closely replicate the coefficient of relative risk aversion of garbage (about 15) using only NIPA consumption (as low as 19). Furthermore, the cross-sectional slope coefficient on 25 Fama-French portfolios is closely replicated using unfiltered NIPA consumption as well (garbage: 2.3%; 2.5% for unfiltered NIPA nondurables & services, and 2.6% for unfiltered NIPA nondurables excluding services), as shown in Table 8. Of course, this modification comes at the cost that I cannot include consumption data for the pre-war period. However, this result shows that, after accounting for filtering, garbage and NIPA consumption behave surprisingly similar.

Table 7: Combining Unfiltering and Q4Q4: GMM Estimates of Relative Risk Aversion

The table reports GMM estimates of the relative risk aversion coefficient γ (RRA). The moment restrictions $g(\gamma)$ are:

$$E\left[\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} R^e_{m,t+1} - 0\right] \text{ and } E\left[\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} R_{f,t+1} - 1\right],$$

where R^e_m is the market excess return and R_f is the gross risk-free rate. The discount parameter is fixed, $\beta = 0.95$. Panel A uses only the first moment restriction (the equity premium), where r_f is the implied log risk-free rate. Panel A uses only the first moment restriction (the equity premium), where r_f is the implied log risk-free rate. Panel B exploits both moment restrictions. The consumption measures are unfiltered fourth-quarter NIPA nondurables & services ($\phi = 0.50$), and fourth-quarter NIPA nondurables excluding services ($\phi = 0.66$).

	Post-War Sample (1950 - 2011)	
	Unfiltered Fourth Quarter NIPA Consumption	
	Nondurables & Serv. $\phi = .50$	Nondurables excl. Serv. $\phi = .66$
Panel A	Equity Premium	
RRA (γ)	22.52	18.98
(s.e.)	(9.76)	(8.51)
Implied r_f, %	27.38	14.40
MAE (g), %	0.00	0.00
Panel A	Equity Premium & R_f	
	One-Stage GMM	
RRA (γ)	45.91	25.45
(s.e.)	(11.79)	(9.23)
MAE (g), %	4.31	1.96
p-value (J)	0.57	0.81
	Two-Stage GMM	
RRA (γ)	8.73	11.13
(s.e.)	(1.78)	(4.42)
MAE (g), %	7.94	5.46
p-value (J)	0.49	0.49

Table 8: Combining Unfiltering and Q4Q4: Fama MacBeth Regressions

The table reports estimates on the consumption risk premium (λ) from the cross-sectional regression of 25 Fama-French portfolios:

$$\bar{R}^e_{i,t} = \lambda_0 + \lambda\beta_i + u_i$$

where $\bar{R}^e_{i,t}$ is mean excess return of portfolio i, λ_0 is a constant (common pricing error), and β_i is the time-series beta of a first pass regression of the portfolio return on a consumption measure. The consumption measures are unfiltered fourth-quarter NIPA nondurables & services ($\phi = 0.50$), and fourth-quarter NIPA nondurables excluding services ($\phi = 0.66$).

	Post-War Sample (1950 - 2011)	
	Unfiltered Fourth Quarter NIPA Consumption	
	Nondurables & Serv. $\phi = .50$	Nondurables excl. Serv. $\phi = .66$
λ_0	-1.93	-1.33
$t_{shanken}$	(-0.44)	(-0.31)
t_{varhac}	[-0.35]	[-0.25]
λ	2.91	2.99
$t_{shanken}$	(2.32)	(2.06)
t_{varhac}	[2.44]	[2.31]
R^2	0.59	0.51
MAE	1.30	1.39
χ^2	(0.018)	(0.019)
	[0.000]	[0.000]
	Estimation without intercept	
λ	2.43	2.63
$t_{shanken}$	(2.95)	(2.86)
t_{varhac}	[2.74]	[2.71]

1.6 Conclusion

Indeed, archaeologists use the trash of our ancestors to apprehend their consumption behavior, and garbage is also useful as a modern consumption measure (Savov (2011)). However, what explains the bad performance of officially reported NIPA consumption? Clever statisticians filter the expenditures of modern humans to estimate consumption. In this paper, I apply a structural model to the NIPA consumption measurement procedure and reconcile three successful alternative consumption measures: Garbage (Savov (2011)), fourth-quarter to fourth-quarter NIPA consumption (Jagannathan and Wang (2007)), and three-year NIPA consumption (Parker and Julliard (2005)).

First of all, filtering explains the relative success of garbage over reported NIPA consumption. This is important, since in the light of the evidence provided it seems unlikely that the success of garbage in measuring consumption is just a coincidence. Furthermore, unfiltered NIPA consumption allows the use of very long time series, a substantial limitation of other alternative consumption measures. Second, filtering intensified by time aggregation provides a possible explanation of the success of NIPA-based alternative consumption measures, specifically three-year and fourth-quarter to fourth-quarter growth. I find that these two measures can be interpreted as ad hoc unfilter rules, where the former is effective in reducing filtering and the latter is effective in reducing time aggregation.

Chapter 2

2 GDP Mimicking Portfolios and the Cross-Section of Stock Returns

This chapter is coauthored by Felix Schindler, Steffen Sebastian, and Erik Theissen.

2.1 Introduction

The Fama and French (1992, 1993) and Carhart (1997) factor models explain the cross-section of asset returns reasonably well. The size, value, and momentum factors used in these models do, however, lack a sound theoretical foundation. A popular alternative, dating back at least to Chen, Roll, and Ross (1986), is to use macroeconomic variables as factors to explain the cross-section of returns. A common approach in this literature is to relate asset returns to changes in GDP or GDP components. GDP in this context can be thought of as a measure of business cycle or recession risk, or as an aggregate which encompasses consumption and investment.

The empirical results are ambiguous. Chen, Roll, and Ross (1986), Campbell (1996) and Griffin and Martin (2003) find that aggregate GDP is not informative for stock returns. Vassalou (2003), on the other hand, shows that a mimicking portfolio designed to capture news on future aggregate GDP growth performs about as well as the Fama-French model. Koijen, Lustig, and Nieuwerburgh (2012) argue that the value premium is a compensation for macroeconomic risk. They show that the Cochrane and Piazzesi (2005)-factor, which forecasts future economic activity (namely the Chicago Fed National Activity Index and aggregate GDP), explains the returns of book-to-market sorted portfolios.

Empirical tests based on GDP components also produce mixed results. The components related to consumption have been extensively analyzed in the literature on consumption-based asset pricing models (see Ludvigson (2012) for a recent survey). Yogo (2006) rationalizes the growth of the stock of durables as a factor. Cochrane (1991, 1996), Li, Vassalou, and Xing (2006), and Belo (2010) advocate the use of investment-related GDP components in asset pricing models.

The empirical success of the models is often judged by how well they explain the returns of 25 portfolios sorted on size and book-to-market.[18] Only few papers have analyzed whether GDP-based models can explain the returns on portfolios sorted by momentum. Again, the results are inconclusive. Chordia and Shivakumar (2002) and Liu and Zhang (2008) conclude that momentum returns are related to macro variables while Griffin and Martin (2003) and Cooper, Gutierrez, and Hameed (2004) find no such relation.

Asset prices should only change when new information becomes available. Against this background the growth rate of aggregate GDP may not be the best choice of a state variable. It is well known that there is a pronounced lead-lag structure in the components of GDP (see e.g. Greenwood and Hercowitz (1991), Gomme, Kydland, and Rupert (2001), Davis and Heathcote (2005), Fisher (2007), Leamer (2007)). In particular, residential investment (RES) leads the GDP, followed by durables (DUR) and nondurables (NDU). Business structures (BST) and equipment and software (EQS), on the other hand, lag GDP. Because aggregate GDP is an average over leading and lagging components, it is a noisy measure of GDP-related news.

In this paper, we investigate the implications of the lead-lag structure of GDP components for the cross-section of asset returns. Ours is the first paper to analyze the explanatory power of aggregate GDP and GDP components in a unified framework. We consider the five GDP components listed above.[19] Our test assets comprise the 25 portfolios sorted on size and book-to-market as well as 10 portfolios sorted on momentum and 30 industry portfolios available from Kenneth French's homepage. We follow Breeden, Gibbons, and Litzenberger (1989), Vassalou (2003), Adrian, Etula, and Muir (2011), among others, and construct factor mimicking portfolios for aggregate GDP and the GDP components.

[18]This practice has been criticized recently by Lewellen, Nagel, and Shanken (2010). They suggest (p. 176) "to include other portfolios in the tests, sorted, for example, by industry, beta, or other characteristics." We follow this suggestion.

[19]We do not consider exports and government spending in this paper. These GDP components play no role in the asset pricing literature. They also do not show a pronounced cyclical pattern (Leamer (2007)).

We find that aggregate GDP does not explain the cross-section of returns. The leading GDP components (residential investment and durables) perform very well in explaining the returns of 25 portfolios sorted on size and book-to-market and do reasonably well in explaining the returns on 10 momentum portfolios. The lagging GDP components, on the other hand, explain the returns of the momentum portfolios very well but fail to explain the returns of the size and book-to-market sorted portfolios. Finally, a three-factor model with the market risk premium, one leading and one lagging GDP component (residential investment and business structures, respectively) compares very favorably with the Carhart four-factor model in jointly explaining the returns on 25 size/book-to-market portfolios, 10 momentum portfolios and 30 industry portfolios.[20]

Our findings have important implications for empirical asset pricing tests using real measures of economic activity as factors. Aggregate GDP is an average of leading and lagging components. However, leading components provide a better explanation of the cross-section of size and book-to-market returns than lagging GDP components or aggregate GDP. Thus, the fact that previous studies did not account for the lead-lag structure may explain why earlier GDP-related asset pricing tests yielded weak results at best. We add further to the literature by showing that lagged GDP components do a good job at explaining the return on momentum portfolios, a result which is in line with Chordia and Shivakumar (2002) and Liu and Zhang (2008). Finally, while some recent papers show that financial variables which predict future aggregate economic activity explain the cross-section of stock returns (e.g. Petkova (2006), Koijen, Lustig, and Nieuwerburgh (2012)), we show that GDP itself also explains the cross-section of returns once the lead-lag pattern of GDP components is taken into account.

The remainder of the paper is structured as follows. In Section 2.2 we discuss the lead-lag structure of GDP components. Section 2.3 describes the construction and characteristics of the factor-mimicking portfolios. In Section **??** we describe our data set and the econometric

[20] In a recent paper Adrian, Etula, and Muir (2011) show that a factor which proxies for shocks to the aggregate leverage of security broker-dealers explains the cross-section of risky asset returns well. Our GDP-based factor model competes favorably with their model. Using a similar test design and a similar set of test assets (but a longer time series), we obtain R^2s of the same order of magnitude as theirs'.

methodology. Section 2.5 presents the results of our asset pricing tests, Section 2.7 concludes.

2.2 Stylized Facts: Lead and Lag in GDP Components

The components of GDP do not move in lockstep, but rather with a quite robust lead and lag to aggregate GDP: residential investment leads the business cycle, followed by durable consumption, and nondurable consumption, followed by the lagging components investment in equipment and software, and investment in business structures. This empirical fact has been neglected in the asset pricing literature, but is well documented in the empirical and theoretical macroeconomic literature.

Empirical Observations. We use annual data for aggregate GDP and components of GDP. Annual data on residential investment, durables, nondurables, equipment and software and business structures come from the Bureau of Economic Analysis (BEA) for the period from 1951 to 2010.[21] Nondurables are measured as "nondurable goods" (NIPA 2.3.4/5, Line 8) and "services" (NIPA 2.3.4/5, Line 13) as is common in the literature. We use the corresponding price indices and population reported by the BEA (NIPA 7.1, Line 18) to compute real per capita growth rates.

The lead and lag pattern in GDP components can best be observed by looking at recessions and recoveries. Recessions materialize from one up to two years earlier in residential investment (the vanguard of the business cycle) than for investment in business structures (the rear guard of the business cycle). Similarly, recoveries can usually be observed first in residential investment and last in business structures.

Figure 4 illustrates the behavior of the average annual per capita real growth rate of each GDP component compared to aggregate GDP during the ten recessions (NBER definition) which occurred between 1951 and 2010. As can be seen, all GDP components closely track the aggregate

[21] NIPA Tables 5.3.4/5, Lines 17, 3, and 9; and NIPA Tables 2.3.4/5, Lines 3, 8, and 19.

from the business cycle peak (zero on the horizontal axis) to the trough and recovery. However, the lead of residential investment and the lag of investment in business structures is easily observable in the figure. Leamer (2007) provides an impressive recession-by-recession comparison of the GDP components. He finds that the lead and lag behavior can be observed not only averaged across all recessions but is a relatively robust feature of each individual recession. Descriptive statistics in Table 9 and forecasting regressions of aggregate GDP on lagged GDP components strongly confirm the lead and lag structure.[22] In the forecasting regressions we regress (using yearly data) the change in per-capita real GDP on the lagged value of the change in the GDP components. Residential investments has by far the largest standardized slope coefficient (0.43) and the largest t-value and R^2 (5.06 and 0.20, respectively). Thus, residential investments are a good predictor of future GDP. The standardized slope of EQS and BST, on the other hand, is negative.

[22] Using the event time approach of Koijen, Lustig, and Nieuwerburgh (2012), we provide further evidence on the lead and lag behavior of GDP components in Appendix B.

Figure 4: Annual GDP Growth During Recessions

The figure shows the average growth rate of aggregate GDP and GDP components during 10 recessions between 1951 and 2010. GDP growth rates are annual, real and on a per capita basis. Vertical lines indicate one standard error confidence intervals. The beginning of recessions correspond to years with the value one on the horizontal axis and exhibit more than two recession quarters according to NBER. Accordingly, year zero is the peak of the cycle, year -1 is one year before, and so forth. In this figure, for a better comparison, the growth rates of aggregate GDP and GDP components are standardized by subtracting the mean and dividing by the standard deviation. Dotted lines in the graphs showing a GDP component correspond to aggregate GDP. The sample period is from 1951 to 2010.

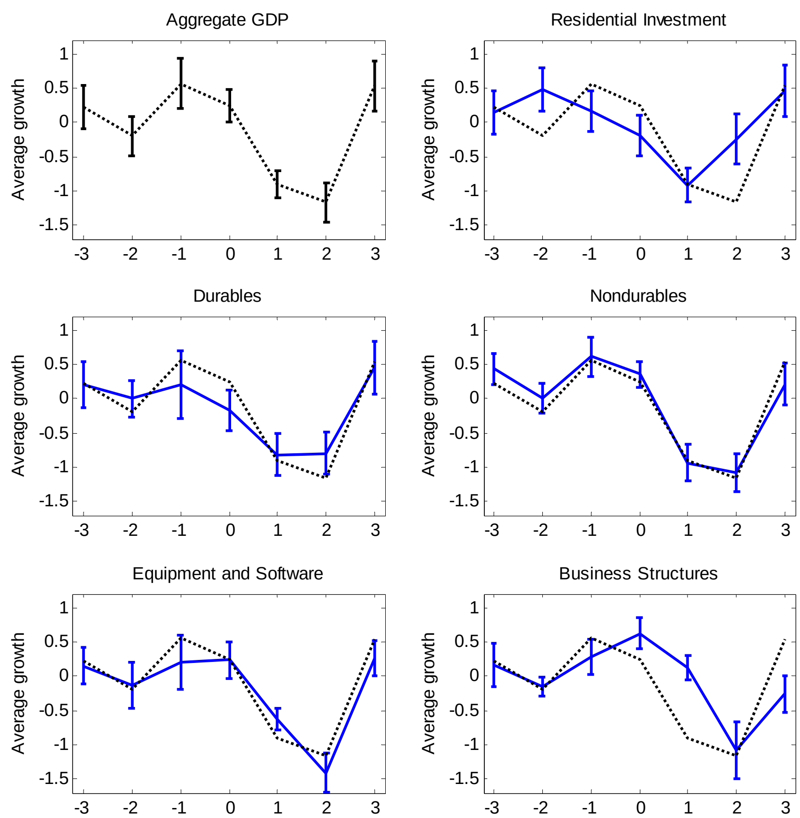


Table 9: Descriptive Statistics

Panel A reports descriptive statistics for U.S. real per capita growth rates for aggregate GDP and components of GDP. MKT is the stock market excess return, as in the Fama/French data library. Panel B provides forecasting regressions of GDP components with respect to future aggregate GDP (all variables are real per capita growth rates). Newey-West corrected t-statistics are reported in parentheses (automatic lag length selection). All data are annual and the sample period is from 1951 to 2010.

Panel A: Leads and lags of real annual per capita growth rates						
			Correlation with			
Aggregate GDP			$\triangle GDP_t$			MKT_t
or GDP Component	Mean	SD	$t-2$	$t-1$	t	t
△Aggregate GDP (GDP)	2.00	2.29	-0.04	0.10	1.00	-0.13
△Residential Investment (RES)	0.57	12.81	0.10	0.44	0.56	0.14
△Durables (DUR)	3.71	6.48	0.09	0.20	0.73	-0.04
△Nondurables (NDU)	1.93	1.27	-0.11	0.30	0.81	-0.01
△Equip. & Software (EQS)	4.46	7.82	-0.03	-0.03	0.82	-0.22
△Business Structures (BST)	0.80	7.53	-0.22	-0.24	0.46	-0.20
Correlation between GDP components			RES	DUR	NDU	EQS
DUR			0.73			
NDU			0.59	0.7		
EQS			0.41	0.71	0.65	
BST			-0.06	0.17	0.43	0.56

Panel B: Forecasting regressions for standardized real per capita growth rates:

$$(\triangle GDP_t - \mu_{GDP}) / \sigma_{GDP} = \alpha + \beta \ (\triangle Y_{j,t-1} - \mu_{Yj}) / \sigma_{Yj} + \epsilon_t$$

	β	$t(\beta)$	R^2
GDP	0.10	(0.72)	0.01
RES	0.43	(4.67)	0.20
DUR	0.20	(1.41)	0.04
NDU	0.30	(2.28)	0.09
EQS	-0.03	(-0.22)	0.00
BST	-0.24	(-1.89)	0.06

Literature. Greenwood and Hercowitz (1991) were the first to theoretically analyze the cyclical behavior of investment in household capital and business capital. Their aim is to construct a model which is able to account for the procyclical behavior of residential investment, and its lead with respect to business investment. The model can generate procyclicality but fails to produce the lead pattern (Greenwood and Hercowitz (1991, p.1210)). Gomme, Kydland, and Rupert (2001) show that including production time brings the model much closer to the data, generating cyclical co-movements and a lag in business investment. Also Davis and Heathcote (2005) calibrate a model which is able to reproduce the fact that residential and nonresidential investment co-move with GDP and consumption. However, their model does not reproduce the empirical fact that "nonresidential investment lags GDP, whereas residential investment leads GDP" (Davis and Heathcote (2005, p.752)). Fisher (2007) provides another possible explanation for why household investment leads nonresidential business investment over the business cycle. He shows that if the traditional home production model is extended such that household capital is complementary to business capital, the model can generate the observed leads and lags.

2.3 GDP-Mimicking Portfolios

GDP and its components are not traded assets, and they are only observable at low frequencies. We therefore follow the literature and construct mimicking portfolios, portfolios of traded assets that track a particular factor.[23] In our analysis of GDP components, we construct tracking portfolios by projecting aggregate GDP and each of the five GDP components discussed in Section 2.2 on the return space of traded assets. We construct six portfolios that have maximum correlation with aggregate GDP and the five GDP components, respectively. These portfolios have the same asset pricing relevant information as the true factors (Balduzzi and Robotti (2008)).[24]

The mimicking portfolio approach has several advantages for our empirical design. First, the returns on stock portfolios can be measured accurately while GDP and its components are observed with error (Breeden, Gibbons, and Litzenberger (1989)).[25] Therefore, statistical

[23]See e.g. Breeden, Gibbons, and Litzenberger (1989), Lamont (2001), Vassalou (2003), Cochrane (2005, p. 109, p. 170), Ferson, Siegel, and Xu (2006), Ang, Hodrick, Xing, and Zhang (2006), Jagannathan and Wang (2007), Balduzzi and Robotti (2008), Cooper and Priestley (2011), Adrian, Etula, and Muir (2011).

[24]The mimicking portfolios are interesting for portfolio advice as well. They can be interpreted as hedges against the risk of the factor they represent (Cochrane (2005, p.167)).

[25]Importantly, it is likely that the different GDP components are subject to different degrees of measurement error, making comparisons between the variables even more difficult.

inference based on mimicking portfolios will be more accurate. Second, stock market returns are observed at higher frequency than GDP and its components. Therefore, we can use monthly data instead of yearly or quarterly data (Breeden, Gibbons, and Litzenberger (1989), Vassalou (2003)). Using monthly data increases the number of time-series observations in our sample. This, in turn, allows us to expand the set of test assets. We include industry and momentum portfolios together with the standard set of 25 size and book-to-market-sorted portfolios. Third, our tests based on mimicking portfolios yield estimates of the risk premia of the factors. These estimates can be compared to the sample mean of the mimicking portfolio excess returns. We are thus able to judge whether the estimated risk premia are plausible, a major concern of Lewellen, Nagel, and Shanken (2010). Finally, when the factors are portfolios of traded assets we can estimate time-series regressions to complement and validate the results of our cross-sectional tests (Jagannathan and Wang (2007)).

In the main paper we only present results obtained using mimicking portfolios. Appendix B.2 provides results of tests based on the real per capita growth rates of aggregate GDP and the five GDP components. The results of these tests are qualitatively similar to those presented in the paper. We therefore conclude that our results are not driven by our mimicking-portfolio-based approach.

Construction. We closely follow Breeden, Gibbons, and Litzenberger (1989), Vassalou (2003) and Adrian, Etula, and Muir (2011) in the construction of our mimicking portfolios. We run the following OLS regression for each of the $j = 1, ..., 6$ annual real per capita growth rates of aggregate GDP and its five components ($\triangle Y_{j,t}$):

$$\triangle Y_{j,t} = a_j + \mathbf{p}_j' \left[\mathbf{FF6}_t,\, WML_t\right] + \epsilon_{j,t}\, \forall j, \tag{14}$$

where $\mathbf{p}_j$ are 7×1 slope coefficients, $\mathbf{FF6}_t$ are annual real excess returns of the six Fama-French portfolios sorted on size and book-to-market (Fama and French (1993)) and WML_t is the momentum factor "winners minus losers" (Carhart (1997)).[26] The selection of these assets is motivated by the well-known fact that they span the mean-variance frontier of a large set of stock market returns (Fama and French (1996)). We normalize the sum of the seven portfolio weights for aggregate GDP and each GDP component to one, $\hat{\mathbf{w}}_j = \hat{\mathbf{p}}_j / \left(\mathbf{1}'\hat{\mathbf{p}}_j\right)^{-1}$. Once we have the mimicking portfolio weights, we can use them to measure the monthly returns of the mimicking portfolios. The mimicking portfolio return for GDP variable j at time t is given by

$$MPm_{j,t} = \hat{\mathbf{w}}_j' \left[\mathbf{FF6m}_t,\, WMLm_t\right], \forall j \tag{15}$$

where $\mathbf{MPm}_t = [MPm_{1,t}, ..., MPm_{6,t}]'$ is a 6×1 vector of mimicking portfolio returns for month t. These returns substitute for the changes in aggregate GDP (GDP), residential investment (RES), durables (DUR), nondurables (NDU), equipment and software (EQS), and business structures (BST).

A Look at the Mimicking Portfolio Factors. The empirical results of the OLS regressions for the construction of the mimicking portfolios are presented in Table 10. The unadjusted R^2 for all OLS regressions are within a range between 10% and 25% and are thus comparable to those in Vassalou (2003) and higher than the ones reported by Lamont (2001). Panel A of the table shows the normalized weights. They already foreshadow some of our main results. The leading GDP components, residential investment and durables, load heavily on the value factor. Consider the mimicking portfolio for residential investments as an example. It has weight 1.63 on small value stocks, weight 0.09 on large value stocks, weight -1.13 on small growth stocks and weight 0.65 on large growth stocks. This results in a total exposure of 2.20 (1.63+0.09-(-1.13)-0.65) to the value factor. The corresponding figure for the mimicking portfolio for durables is 2.00. By contrast, the GDP components lagging the business cycles such as equipment and software and business structures have much lower exposure to value (1.21 and 0.84, respectively). The pattern for the exposure to size is quite similar. Residential investments and durables have the

[26] We use stock return data available from the web site of Kenneth R. French (Fama and French (1993)). See Section **??** for details on the data.

highest size exposure (0.47=-1.13-0.01+1.63-0.65-(-0.72)-0.09) and 0.79=-1.92+0.40+2.00-1.39-(-1.17)-(-0.53), respectively) while business structure has the lowest size exposure (-1.56=0.23-0.25-0.92-(-0.95)-0.53-1.04).

This picture changes completely when we consider the exposure to momentum in the last row of Panel A in Table 10. Here, the mimicking portfolios for residential investment, nondurables, and durables have the lowest weights while the mimicking portfolios for the GDP components lagging the business cycle are highly invested in the momentum portfolio.

We further analyze the mean-variance properties of the mimicking portfolios (similar to Jagannathan and Wang (2007) and Adrian, Etula, and Muir (2011)). The first line of Panel B of Table 10 shows the mean return of the mimicking portfolios. As mentioned above we will use these mean returns as benchmark values for the factor risk premia that we estimate in our asset pricing tests. Lines 2 and 3 of Panel B show the return standard deviation and the Sharpe ratios of the mimicking portfolios. It is noteworthy that the Sharpe ratios decline monotonically as we move from the leading GDP components to the lagging components. Figure 5 plots the efficient frontier based on the six Fama-French portfolios and the momentum portfolio. It further shows the mimicking portfolios for each of the five GDP components and the one for aggregate GDP. The mimicking portfolio for residential investments is reasonably close to the efficient frontier (although the GRS statistic rejects the null of efficiency).

The lower part of Panel B shows the intersect and slope estimator of time series regressions in which we regress the returns of the GDP-mimicking portfolios on the four factors of the Carhart model. The coefficients are mainly statistically significant and match with the findings from Panel A of Table 10. The exposure to the size and value factor declines as we move from the leading to the lagging GDP components while the reverse is true for the momentum factor. It is also interesting to note that the market beta is positive for the leading GDP components and negative for the lagging components.[27]

[27] This correlation pattern of GDP components and stock returns also shows up visually when we follow the event time methodology proposed by Koijen, Lustig, and Nieuwerburgh (2012), as we show in Appendix B.

Table 10: GDP Component Mimicking Portfolios

Panel A reports estimates of $j = 1, ..., 6$ mimicking portfolio weights for aggregate GDP (GDP), residential investment (RES), durables (DUR), nondurables (NDU), equipment and software (EQS) and business structures (BST). We regress the real per capita growth rate of annual aggregate GDP and each of the annual GDP components ($\triangle Y_{j,t}$) on six annual Fama-French portfolios, $\mathbf{FF6}_t$, sorted by size and book-to-market (low/small, mid/small, high/small, low/big, mid/big, high/big) and the momentum factor, WML_t:

$$\triangle Y_{j,t} = a_j + \mathbf{p}_j' \left[\mathbf{FF6}_t,\ WML_t\right] + \epsilon_{j,t}.$$

The resulting weights $\hat{\mathbf{p}}_j$ are normalized, such that their sum is one: $\hat{\mathbf{w}}_j = \hat{\mathbf{p}}_j \left(\mathbf{1}'\hat{\mathbf{p}}_{,j}\right)^{-1}$. Monthly GDP mimicking portfolios are calculated as: $MPm_{j,t} = \hat{\mathbf{w}}_j' \left[\mathbf{FF6m}_t,\ WMLm_t\right]$, where $\mathbf{FF6m}_t$, and $WMLm_t$ are monthly measured returns of the six Fama-French portfolios and the momentum factor. Panel B reports the mean (%), standard deviation (%), Sharpe ratio and the factor exposures (betas) of the six monthly GDP mimicking portfolios:

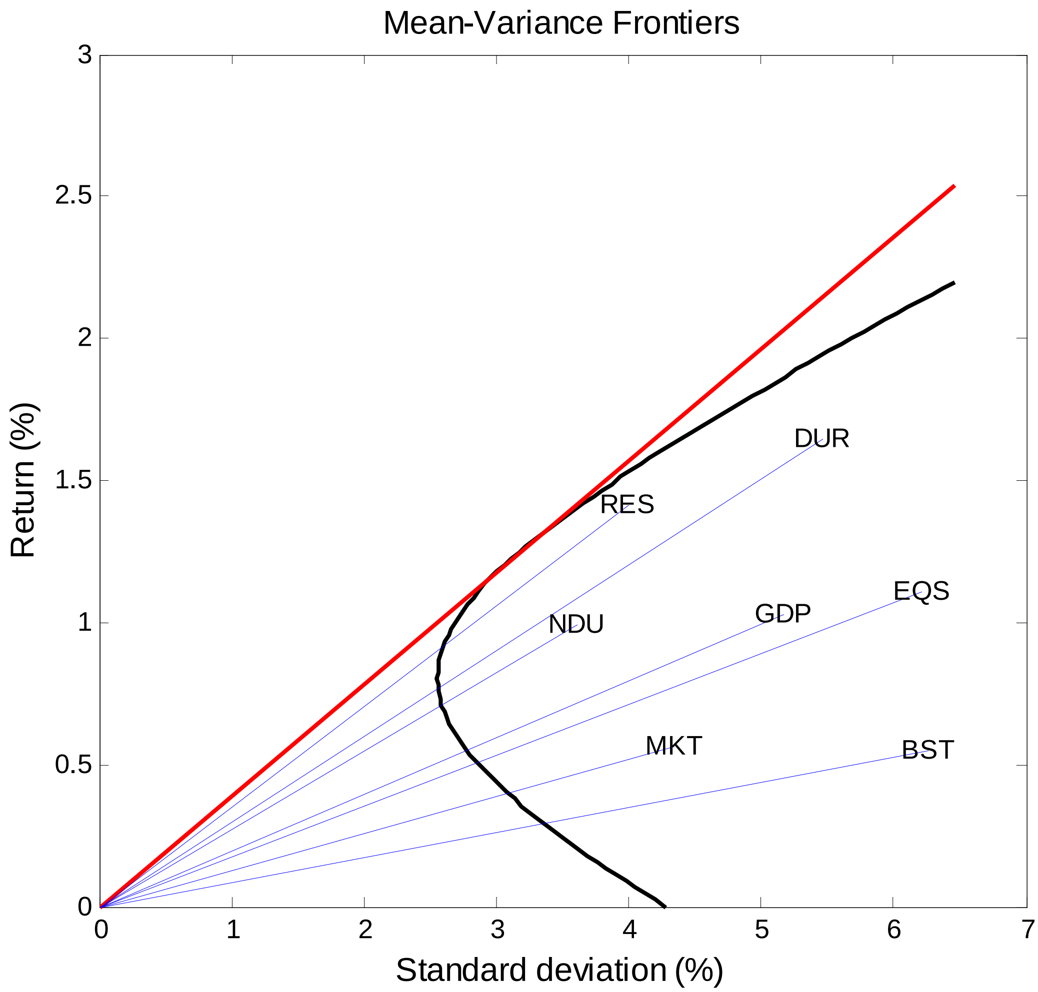
$$MPm_{j,t} = \alpha_j + \beta_{MKT,j} MKT_t + \beta_{SMB,j} SMB_t + \beta_{HML,j} HML_t + \beta_{WML,j} WML_t,$$

where MKT is the market excess return, SMB is the small minus big factor, HML is the high minus low factor, and WML is the winner minus loser (momentum) factor. The sample period is from January 1951 to December 2010.

	GDP	RES	DUR	NDU	EQS	BST
Panel A:	Normalized GDP mimicking portfolio weights					
low/small	-0.79	-1.13	-1.92	-0.48	-0.75	0.23
mid/small	-0.30	-0.01	0.40	0.01	-1.37	-0.25
high/small	0.83	1.63	2.00	0.31	1.97	-0.92
low/big	0.04	0.65	1.39	-0.24	0.57	-0.95
mid/big	0.52	-0.72	-1.17	0.04	0.27	0.53
high/big	-0.39	0.09	-0.53	0.56	-0.94	1.04
momentum	1.09	0.49	0.84	0.80	1.25	1.33
Panel B:	Statistics of monthly mimicking portfolio factors					
Mean (%)	1.03	1.41	1.65	0.99	1.11	0.55
SD (%)	5.17	4.00	5.47	3.61	6.22	6.27
Sharpe ratio	0.20	0.35	0.30	0.28	0.18	0.09
α	0.00	0.00	0.01	0.00	0.00	-0.00
s.e.	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
β_{MKT}	-0.17	0.46	0.02	0.20	-0.30	-0.25
s.e.	(0.01)	(0.02)	(0.04)	(0.00)	(0.03)	(0.02)
β_{SMB}	-0.41	0.23	0.02	-0.18	-0.31	-0.69
s.e.	(0.02)	(0.04)	(0.07)	(0.00)	(0.04)	(0.03)
β_{HML}	0.54	1.09	0.88	0.85	0.30	0.43
s.e.	(0.02)	(0.04)	(0.07)	(0.00)	(0.05)	(0.03)
β_{WML}	1.12	0.53	0.92	0.80	1.30	1.29
s.e.	(0.02)	(0.03)	(0.05)	(0.00)	(0.03)	(0.02)

Figure 5: Mean-Variance Frontier and GDP Mimicking Portfolio Factors

The figure shows the mean-variance frontier of six Fama-French portfolios sorted by size and book-to-market and the momentum factor. RES is the mimicking portfolio of residential investment. DUR, NDU, EQS and BST are the analogous mimicking portfolios for durables, nondurables, equipment and software and business structures. The mimicking portfolios are based on the same assets which are represented by the mean-variance frontier. MKT is the market excess return. All data are monthly and the sample period is from January 1951 to December 2010.



2.4 Model, Estimation, and Data

Model. Recessions are bad news for investors (Chen, Roll, and Ross (1986), Campbell (2000), Cochrane (2005), p.172): The prospect of an upcoming recession gives reason to expect a lower level of production, lower dividends and higher distress risk for equity investments. Investors with jobs are subject to larger idiosyncratic labor income and unemployment risk than in normal times. In short, stochastic discount factors should be highest at the onset of recessions. Thus, investors avoid assets which perform poorly in recessions, and such stocks should compensate investors with larger expected returns.

We use the stochastic discount factor (SDF) representation to estimate linear factor models. More precisely, we consider the empirical SDF specification

$$m_t = 1 - \widetilde{\mathbf{f}}_t^{\prime}\mathbf{b}, \tag{16}$$

where m_t denotes the SDF, $\widetilde{\mathbf{f}}_t$ are K de-meaned factors, $\widetilde{\mathbf{f}}_t = (\mathbf{f}_t - \boldsymbol{\mu})$, and $\mathbf{b}$ are K factor loadings. We use an ICAPM specifications and therefore include the market excess return, MKT_t as a factor (Fama (1996)). In most specifications, we augment the SDF with *one* of the j GDP component mimicking portfolios. It is included as a factor capturing recession risk.

The linear SDF specification in Equation (16) implies that N excess returns, $\mathbf{R}_t$, are related to the factors by (Burnside (2011)):

$$E\left(\mathbf{R}_t\right) = cov\left(\mathbf{R}_t,\ \mathbf{f}_t\right)\mathbf{b}. \tag{17}$$

A given estimate of $\mathbf{b}$ allows to make an inference on whether a specific factor helps to price the considered set of assets given the other factors (Cochrane (2005)). Equation (17) underscores also an economic restriction on the sign of the estimated SDF loadings $\mathbf{b}$. Theory suggests that an economically sensible SDF should take on high values in “bad times” (recessions) and low values in “good times”. Assets which covary counter-cyclically with the SDF should provide higher expected returns (Campbell (2000), Cochrane (2005)). All five components of GDP are pro-cyclical (see Section 2.2). Therefore, the estimates of the SDF loading for a specific GDP component (and its mimicking portfolio) should be positive.[28]

The SDF representation is transferable to the traditional expected return-beta representation. Rearranging Equation (17) gives:

$$E(\mathbf{R}_t) = \boldsymbol{\beta}\boldsymbol{\lambda}, \tag{18}$$

where $\boldsymbol{\beta} = cov(\mathbf{R}_t, \mathbf{f}_t)\,\boldsymbol{\Sigma}_{ff}^{-1}$ is a $N \times K$ matrix of factor betas, $\boldsymbol{\lambda} = \boldsymbol{\Sigma}_{ff}\mathbf{b}$ is a $K \times 1$ vector of factor risk premia, and $\boldsymbol{\Sigma}_{ff}$ is the covariance matrix of factors. Given a positive factor risk premium, a specific asset should offer higher expected returns if its factor beta is positive. The factor risk premium can be interpreted as the price the SDF assigns to the factor. If the factor is a traded asset (e.g., a GDP mimicking portfolio), the estimated factor risk premium should be equal to its sample mean (and the factor should thus price itself). We exploit this relation in order to check wether the estimated factor risk premia are economically plausible (Lewellen, Nagel, and Shanken (2010)).

Estimation. Our estimation methodology follows Cochrane (2005) and Burnside (2010, 2011). We report heteroscedasticity and autocorrelation-robust Generalized Method of Moments (GMM) estimates for SDF loadings $\mathbf{b}$ and implied factor risk premia $\boldsymbol{\lambda}$. In principle, if the model is specified correctly, several normalizations of the SDF are equivalent. However, Burnside (2010) shows that the SDF specification with de-meaned factors has greater power to reject misspecified models than other normalizations. We include a common pricing error (or constant) in all models. This constant should be zero for each model. Testing the estimated constant against zero allows us to check the validity of the model. We re-estimated all models without including a constant. The results are shown in Appendix B.

We estimate linear factor models using the $N + K + \kappa$ empirical moment restrictions

[28] Recently, Maio and Santa-Clara (2012) have shown that many popular multifactor models do not meet this criteria. Estimated SDF loadings frequently have the wrong sign and are thus inconsistent with the ICAPM.

$$\mathbf{g} = g_T\left[\gamma,\ \mathbf{b},\ \boldsymbol{\mu},\ vec\left(\boldsymbol{\Sigma}_{ff}\right)\right] = E_T \begin{bmatrix} \mathbf{R}_t\left[1 - \left(\mathbf{f}_t - \boldsymbol{\mu}\right)' \mathbf{b}\right] - \gamma \\ \mathbf{f}_t - \boldsymbol{\mu} \\ vec\left[\left(\mathbf{f}_t - \boldsymbol{\mu}\right)\left(\mathbf{f}_t - \boldsymbol{\mu}\right)'\right] - vec\left(\boldsymbol{\Sigma}_{ff}\right) \end{bmatrix}, \tag{19}$$

where γ is a common pricing error (a constant) and $vec\left(\boldsymbol{\Sigma}_{ff}\right)$ denotes the $\kappa = K\left(K+1\right)/2$ unique elements of the covariance matrix of factors. We use first-stage GMM with the standard identity matrix as a weighting matrix and require the GMM estimator to set $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}_{ff}$ equal to their sample counterparts. Burnside (2011) shows that using the moment conditions in Equation (19) along with the identity weighting matrix results in numerical equivalence between the SDF-based GMM estimates and the two-pass Fama-MacBeth estimates of γ and $\boldsymbol{\lambda}$ (Fama and MacBeth (1973)). This fact makes our estimates easy to interpret and comparable to the vast literature that uses the Fama-MacBeth two-stage estimation procedure. Standard errors of the parameters are based on the HAC-robust covariance estimator of $\mathbf{S} = \sum_{j=-\infty}^{\infty} E\left[\mathbf{g}\mathbf{g}'\right]$ proposed by Newey and West (1987) using the automatic lag length selection procedure of Andrews (1991). As described in Cochrane (2005) and Burnside (2010, 2011), adding the additional $K + \kappa$ moment conditions facilitates the correction of the standard errors of $\hat{\gamma}$, $\hat{\mathbf{b}}$ and $\hat{\boldsymbol{\lambda}}$ for the fact that $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}_{ff}$ are estimated. As measures of model fit we report the cross-sectional OLS R^2, the mean absolute error (MAE), and the Hansen-Jagannathan distance (HJ) with the simulation-based p-value (Hansen and Jagannathan (1997), Kan and Zhou (2012a)). We also report the cross-sectional GLS R^2. It measures how close the best combination of the factors is to the mean-variance frontier of tested assets and serves to illustrate the economic validity of the cross-sectional fit (Kandel and Stambaugh (1995), Lewellen, Nagel, and Shanken (2010)).

Return Data. For the asset pricing tests, we use the 25 Fama-French portfolios sorted by size and book-to-market, 10 momentum portfolios, 30 industry portfolios, the market premium, SMB, HML, WML (the momentum factor), and the one-period T-bill rate available on the web site of Kenneth R. French (Fama and French (1993)). The momentum returns use the "2-12" convention, i.e. they are based on the returns from the previous 12 month excluding the last one.

We use annual data from 1951 to 2010, and monthly data from January 1951 to December 2010; stock market returns are in excess of the one-period T-bill rate. We deflate yearly returns by the price index implied by our consumption measure (nondurables). To deflate monthly returns, we use the monthly CPI available from the Bureau of Labor Statistics.

2.5 Asset Pricing Tests

2.5.1 Dissecting GDP and the Cross-Section of Stock Returns

Size and Value. We use monthly GDP component mimicking portfolios throughout this section. Table 11 shows our test results when we use 25 size and book-to-market-sorted portfolios as test assets. We see from specification (1) that the market factor is unable to explain the average returns of the size and book-to-market portfolios. The constant (the common pricing error) is large (1.34%) and significant; the SDF loading b as well as the factor risk premium λ are negative and insignificant.[29] Adding aggregate GDP as a factor does not improve the model (specification (2)) .

Next, we substitute aggregate GDP with *one* of the GDP components in the same order in which the GDP components lead and lag aggregate GDP. In specification (3), we include residential investment, which is the vanguard of the business cycle. The constant is relatively small, 0.30%, and insignificant. The SDF loading on residential investment and the corresponding factor risk premium are positive and significant at conventional levels. The OLS (GLS) R^2 is about 0.76 (0.46), and the HJ-distance is smaller than in specification (1) and (2). However, there is also some evidence of misspecification. The estimated factor risk premium for residential investment is only 1.11%, as compared to the sample mean of the factor of 1.41%, and the factor price for the market excess return is too small as well. In specification (4), we find a similar model fit for durables. The constant is economically and statistically marginal, 0.05%, and the SDF loading is positive and significant. The model explains 72% of the cross-sectional variation

[29] All factors are traded assets. Thus, we can interpret the pricing error economically as an annualized return of 12×1.34%=16.08% not explained by the factors. Note that the estimated constant and the factor prices are numerically equivalent to the traditional Fama-MacBeth estimates (Burnside (2011)).

in size and book-to-market sorted portfolio returns (with a GLS R^2 of 0.43). The next row, specification (5), shows results for nondurables. The SDF loading and the factor price are still positive and significant. However, all measures of model fit indicate a less favorable model fit, e.g. the OLS (GLS) R^2 is only 0.26 (0.16).

In line (6), the table reports results for equipment and software. We find a large and significant constant of 1.65%. The SDF loading is negative (insignificant), the estimate of the factor risk premium is negative (insignificant) and the measures of model fit are similar to those for nondurables. Since all GDP components are pro-cyclical, the SDF loading should be positive. We conclude that the estimate in specification (6) has the wrong sign. This is confirmed when comparing the estimate of the factor risk premium to the mean of the factor: The estimated risk premium is -0.82% and is more than two standard errors away from the sample mean of +1.11% (see Table 10). In specification (7), we find very similar results for business structures. Even though the measures of model fit suggest a good fit (OLS (GLS) R^2 of 0.68 (0.31)), the model is clearly misspecified since the constant is large (2.39%) and the slope coefficient is negative and thus economically implausible.

In summary, we find that the leading GDP components are able to explain the returns of 25 size and book-to-market-sorted portfolios reasonably well. The coefficients have the expected signs, are significant, and explain up to 76% of the cross-sectional variation of the returns on the test assets. The lagging GDP components, on the other hand, yield negative estimates of the factor risk premium. Appendix B.1 shows the estimates of restricted models in which we force the constant to be zero. We find that our results are robust to this alternative specification. In Appendix B.2 we report results based on actual (annual) GDP data rather than on mimicking portfolios. They are qualitatively similar.

Table 11: GDP Mimicking Portfolio Factors: 25 Fama-French Portfolios

The table reports GMM estimates of asset pricing models given by

$$E(\mathbf{R}_t) = \gamma + \boldsymbol{\beta}\boldsymbol{\lambda}.$$

Estimates of the SDF loadings $\mathbf{b}$ and a constant γ are obtained by exploiting the moment restrictions $E(\mathbf{R}_t[1 - (\mathbf{f}_t - \boldsymbol{\mu})'\mathbf{b}] - \gamma) = 0$ and $E(\mathbf{f}_t - \boldsymbol{\mu}) = 0$, where $\mathbf{R}_t$ is a vector of stock returns and $\mathbf{f}_t$ is a vector of risk factors. The loadings on the stochastic discount factor are multiplied by 100 in the table. Factor risk prices are calculated as $\boldsymbol{\lambda} = \boldsymbol{\Sigma}_{ff}\mathbf{b}$, standard errors are obtained by the delta method, and are corrected for the fact that the covariance matrix of risk factors $\boldsymbol{\Sigma}_{ff}$ is estimated. The stock returns are 25 value-weighted Fama-French portfolios sorted by size and book-to-market. The risk factors are the excess return on the market portfolio (MKT), and factor mimicking portfolios for aggregate GDP (GDP), residential investment (RES), durables (DUR), nondurables (NDU), equipment and software (EQS), and business structures (BST). Estimation is by first-stage GMM using the identity matrix. T-statistics are reported in parentheses and are based on the Newey and West (1987) HAC covariance matrix with automatic lag length selection according to Andrews (1991). We report the cross-sectional OLS R^2, below the cross-sectional GLS R^2, the mean absolute error (MAE), and the Hansen-Jagannathan distance (HJ) with the simulation-based p-value for a test of a zero HJ. All data are monthly and the sample period is from January 1951 to December 2010.

	Const	SDF Loadings ($\mathbf{b}$)		Factor Prices ($\boldsymbol{\lambda}$)		$R^{2\,OLS}_{GLS}$	MAE	HJ(pv)
(1)	γ	MKT		MKT				
	1.34	-2.91		-0.55		0.13	0.15	0.28
	(3.41)	(-1.28)		(-1.27)		0.10		(0.00)
(2)	γ	MKT	GDP	MKT	GDP			
	2.47	-12.61	-7.31	-1.76	-0.89	0.26	0.13	0.28
	(3.47)	(-2.28)	(-1.59)	(-2.50)	(-1.05)	0.14		(0.00)
(3)	γ	MKT	RES	MKT	RES			
	0.30	-0.24	7.06	0.32	1.11	0.76	0.07	0.22
	(0.67)	(-0.09)	(3.76)	(0.63)	(3.10)	0.46		(0.07)
(4)	γ	MKT	DUR	MKT	DUR			
	0.05	5.49	8.56	0.68	2.33	0.72	0.08	0.23
	(0.10)	(1.53)	(4.29)	(1.19)	(4.52)	0.43		(0.04)
(5)	γ	MKT	NDU	MKT	NDU			
	0.12	3.68	7.54	0.61	0.94	0.26	0.14	0.28
	(0.29)	(1.25)	(2.24)	(1.24)	(2.26)	0.16		(0.00)
(6)	γ	MKT	EQS	MKT	EQS			
	1.65	-7.15	-4.02	-0.93	-0.82	0.20	0.14	0.27
	(3.29)	(-1.98)	(-1.52)	(-1.81)	(-1.00)	0.19		(0.00)
(7)	γ	MKT	BST	MKT	BST			
	2.39	-13.49	-7.27	-1.74	-1.38	0.68	0.09	0.25
	(4.13)	(-3.34)	(-2.84)	(-3.34)	(-1.82)	0.31		(0.04)

Momentum. Table 12 shows our test results when we use 10 momentum-sorted portfolios as test assets. The structure of the table is similar to Table 11. Specification (1) shows the traditional CAPM, specification (2) a model with aggregate GDP, and specifications (3) to (7) show the performance of the GDP components as risk factors beginning with the vanguard GDP component (residential investment) and closing with the rear guard of the business cycle (business structures).

The leading GDP components continue to perform well. They produce small pricing errors (i.e., a small and insignificant constant), positive and significant factor risk premia, and reasonably high R^2s. Once we consider the lagging GDP components, however, we find that our previous results are turned upside down. We find that the lagging GDP components capture the average returns of the momentum portfolios quite well, and better than the leading GDP components. Focusing on the rear guard of the business cycle, business structures in specification (7), the constant is small (0.36%), the SDF loading is *positive* and significant at conventional levels. This stands in stark contrast to the result based on size and book-to-market-sorted portfolios where we found *negative* estimates of the SDF loading and factor risk premia. However, there are also indications of misspecification. In particular, the estimated factor risk premium (1.06%), albeit having the right sign, is too large compared to the sample mean (0.55%).

Appendix B.1 presents results of restricted models in which we force the constant to be zero. Again, we find that our results are robust to this alternative specification. Appendix B.2 shows results based on actual (annual) GDP data rather than on mimicking portfolios. They are qualitatively similar. However, the leading GDP components underperform the lagging components more clearly.

Table 12: GDP Mimicking Portfolio Factors: 10 Momentum Portfolios

The table reports GMM estimates of asset pricing models given by

$$E(\mathbf{R}_t) = \gamma + \boldsymbol{\beta}\boldsymbol{\lambda}.$$

Estimates of the SDF loadings $\mathbf{b}$ and a constant γ are obtained by exploiting the moment restrictions $E(\mathbf{R}_t[1 - (\mathbf{f}_t - \boldsymbol{\mu})'\mathbf{b}] - \gamma) = 0$ and $E(\mathbf{f}_t - \boldsymbol{\mu}) = 0$, where $\mathbf{R}_t$ is a vector of stock returns and $\mathbf{f}_t$ is a vector of risk factors. The loadings on the stochastic discount factor are multiplied by 100 in the table. Factor risk prices are calculated as $\boldsymbol{\lambda} = \boldsymbol{\Sigma}_{ff}\mathbf{b}$, standard errors are obtained by the delta method, and are corrected for the fact that the covariance matrix of risk factors $\boldsymbol{\Sigma}_{ff}$ is estimated. The stock returns are 10 value-weighted momentum portfolios (2-12). The risk factors are the excess return on the market portfolio (MKT), and factor mimicking portfolios for aggregate GDP (GDP), residential investment (RES), durables (DUR), nondurables (NDU), equipment and software (EQS), and business structures (BST). Estimation is by first-stage GMM using the identity matrix. T-statistics are reported in parentheses and are based on the Newey and West (1987) HAC covariance matrix with automatic lag length selection according to Andrews (1991). We report the cross-sectional OLS R^2, below the cross-sectional GLS R^2, the mean absolute error (MAE), and the Hansen-Jagannathan distance (HJ) with the simulation-based p-value for a test of a zero HJ. All data are monthly and the sample period is from January 1951 to December 2010.

	Const	SDF Loadings ($\mathbf{b}$)		Factor Prices ($\boldsymbol{\lambda}$)		$R^{2\,OLS}_{GLS}$	MAE	HJ(pv)
(1)	γ	MKT		MKT				
	1.52	-4.96		-0.93		0.18	0.23	0.23
	(4.27)	(-2.20)		(-2.20)		0.01		(0.00)
(2)	γ	MKT	GDP	MKT	GDP			
	0.25	4.04	4.43	0.39	0.84	0.89	0.10	0.18
	(0.74)	(1.70)	(3.17)	(0.99)	(3.50)	0.39		(0.01)
(3)	γ	MKT	RES	MKT	RES			
	-0.06	-0.05	12.33	0.62	1.96	0.86	0.11	0.20
	(-0.14)	(-0.02)	(2.89)	(1.49)	(3.25)	0.27		(0.00)
(4)	γ	MKT	DUR	MKT	DUR			
	0.17	3.70	5.87	0.46	1.60	0.89	0.10	0.19
	(0.46)	(1.59)	(3.11)	(1.15)	(3.71)	0.37		(0.01)
(5)	γ	MKT	NDU	MKT	NDU			
	0.16	2.77	6.74	0.45	0.85	0.88	0.10	0.19
	(0.49)	(1.29)	(3.17)	(1.18)	(4.16)	0.35		(0.00)
(6)	γ	MKT	EQS	MKT	EQS			
	0.43	3.13	3.62	0.22	1.08	0.90	0.10	0.18
	(1.31)	(1.38)	(3.13)	(0.57)	(3.54)	0.41		(0.01)
(7)	γ	MKT	BST	MKT	BST			
	0.36	3.69	3.72	0.29	1.06	0.89	0.10	0.18
	(1.11)	(1.61)	(3.14)	(1.61)	(3.31)	0.40		(0.01)

2.5.2 A GDP-Based Three-Factor Model

From our previous analysis we know that leading GDP components can explain the average returns of size and book-to-market portfolios and lagging GDP components can explain the average returns of momentum portfolios. Thus, when we test size, value, and momentum portfolios jointly, it is natural to consider a GDP-based three-factor model which includes the market excess return, residential investment as the most leading GDP component, and business structures as the most lagging GDP component. This specification should captures the lead and lag structure of the GDP components. The two GDP components are chosen to share the job of pricing a large cross-section of assets: The task of residential investments is to price the size and book-to-market-sorted portfolios while the task of business structures is to price the momentum portfolios.

Size, Value, and Momentum. Table 13 shows the results that we obtain when we simultaneously use 25 double sorted size/book-to-market portfolios and 10 single sorted momentum portfolios as test assets. In specifications (1) and (2) we find that the traditional CAPM and the CAPM augmented by aggregate GDP as a state variable are not able to explain the returns of the test assets. Scrolling through the two-factor specifications (3) to (7), the table shows that leading GDP components perform better in pricing the 35 portfolios than the lagging GDP components.

Results for our GDP-based three-factor model are reported in line (8). The estimated constant is very close to zero (0.01%) and insignificant (t-statistic of 0.03). The estimated factor risk premium for residential investment is 1.44% (t: 4.40), which is indistinguishable from the sample mean of 1.41%. Similarly, the estimated risk premium for business structures, 0.54% (t: 1.81), is significant and very close to its sample mean of 0.55%. Even the estimated market risk premium, 0.59%, is indistinguishable from its sample mean of 0.56% (t: 1.24). Thus, all four point estimates, the constant and the three factor risk premia, are economically sensible. The OLS (GLS) R^2 is 0.82 (0.42), suggesting a good model fit. The good fit is confirmed by a low

MAE of 0.09. Only the HJ-distance statistic has not improved much as compared to the other specifications, although it does not reject the three-factor model at the 1% level.

Figure 6 provides a graphical illustration of the 35 individual pricing errors. The CAPM fails to price the 35 test assets. Adding business structures (the same holds for aggregate GDP) does not help to reduce the pricing errors of the 25 size/book-to-market portfolios. Adding the vanguard GDP component, residential investment, instead results in a much better fit. The size/book-to-market portfolios move towards the diagonal line. Including both residential investment and business structure results in even lower pricing errors. Most importantly, the extreme momentum portfolios (MomH and MomL) move closer to the 45° line.

Again, Appendix B.1 shows the results that we obtain when we impose a zero constant. They are qualitatively similar to the results presented in the text. Appendix B.2 shows results based on actual (annual) GDP data rather than on mimicking portfolios. The three factor model performs well. We estimate significantly positive factor risk premia for the market excess return and residential investments. The R^2 is 0.63. The risk premium for business structures is positive but insignificant.

Table 13: GDP Mimicking Factors: 25 Fama-French and 10 Momentum Portfolios

The table reports GMM estimates of asset pricing models given by

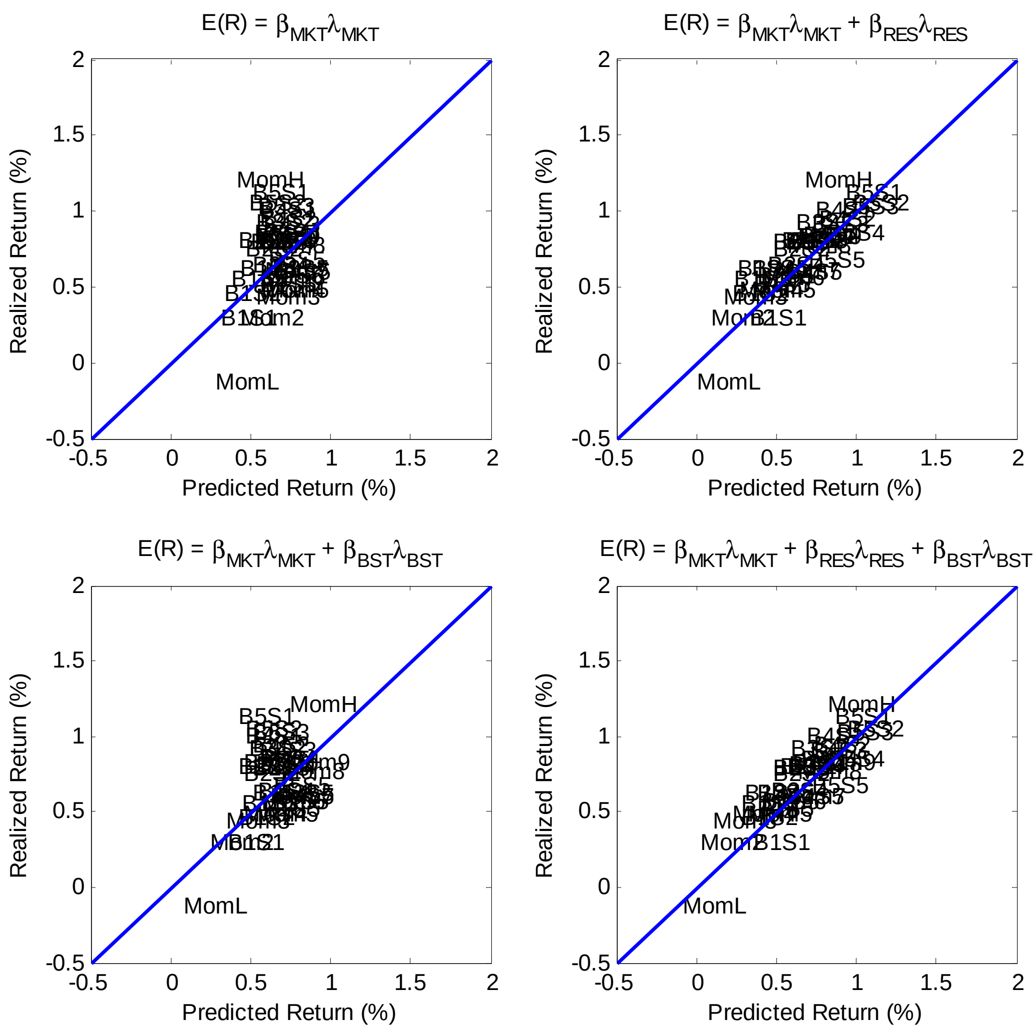
$$E(\mathbf{R}_t) = \gamma + \boldsymbol{\beta}\boldsymbol{\lambda}.$$

Estimates of the SDF loadings $\mathbf{b}$ and a constant γ are obtained by exploiting the moment restrictions $E(\mathbf{R}_t[1 - (\mathbf{f}_t - \boldsymbol{\mu})'\mathbf{b}] - \gamma) = 0$ and $E(\mathbf{f}_t - \boldsymbol{\mu}) = 0$, where $\mathbf{R}_t$ is a vector of stock returns and $\mathbf{f}_t$ is a vector of risk factors. The loadings on the stochastic discount factor are multiplied by 100 in the table. Factor risk prices are calculated as $\boldsymbol{\lambda} = \boldsymbol{\Sigma}_{ff}\mathbf{b}$, standard errors are obtained by the delta method, and are corrected for the fact that the covariance matrix of risk factors $\boldsymbol{\Sigma}_{ff}$ is estimated. The stock returns are 25 Fama-French portfolios sorted by size and book-to-market and 10 momentum portfolios (2-12). Estimation is by first-stage GMM using the identity matrix. T-statistics are reported in parentheses and are based on the Newey and West (1987) HAC covariance matrix with automatic lag length selection according to Andrews (1991). We report the cross-sectional OLS R^2, below the cross-sectional GLS R^2, the mean absolute error (MAE), and the Hansen-Jagannathan distance (HJ) with the simulation-based p-value for a test of a zero HJ. All data are monthly and the sample period is from January 1951 to December 2010.

	Const	SDF Loadings ($\mathbf{b}$)			Factor Prices ($\boldsymbol{\lambda}$)			R^{2OLS}_{GLS}	MAE	HJ(pv)
(1)	γ	MKT			MKT					
	1.35	-3.32			-0.62			0.11	0.20	0.36
	(4.23)	(-1.67)			(-1.65)			0.04		(0.00)
(2)	γ	MKT	GDP		MKT	GDP				
	0.57	2.85	3.96		0.20	0.82		0.41	0.18	0.33
	(1.43)	(1.06)	(2.80)		(0.47)	(3.80)		0.24		(0.00)
(3)	γ	MKT	RES		MKT	RES				
	0.16	-0.18	8.80		0.42	1.39		0.79	0.09	0.29
	(0.38)	(-0.07)	(4.79)		(0.88)	(4.25)		0.42		(0.02)
(4)	γ	MKT	DUR		MKT	DUR				
	0.16	4.41	6.86		0.55	1.87		0.78	0.10	0.29
	(0.35)	(1.50)	(3.97)		(1.12)	(5.37)		0.41		(0.02)
(5)	γ	MKT	NDU		MKT	NDU				
	0.08	3.73	7.43		0.62	0.93		0.60	0.14	0.33
	(0.19)	(1.47)	(3.68)		(1.39)	(5.19)		0.25		(0.00)
(6)	γ	MKT	EQS		MKT	EQS				
	0.87	1.29	3.19		-0.08	1.10		0.42	0.18	0.32
	(2.40)	(0.52)	(2.67)		(-0.20)	(3.86)		0.27		(0.00)
(7)	γ	MKT	BST		MKT	BST				
	0.92	0.47	2.17		-0.15	0.80		0.25	0.20	0.36
	(2.59)	(0.20)	(2.10)		(-0.37)	(2.88)		0.08		(0.00)
(8)	γ	MKT	RES	BST	MKT	RES	BST			
	0.01	1.50	8.32	1.06	0.59	1.44	0.54	0.82	0.09	0.29
	(0.03)	(0.56)	(4.45)	(0.97)	(1.24)	(4.40)	(1.81)	0.42		(0.01)

Figure 6: Predicted versus Realized Returns

The figure shows pricing errors for 25 Fama-French portfolios sorted by size and book-to-market and 10 momentum portfolios (2-12) using monthly GDP mimicking portfolios. GDP is aggregate GDP, RES is residential investment, BST is business structures, and MKT is the market excess return. All data are monthly and the sample period is from January 1951 to December 2010.



Time-Series Regressions. It is a well-established fact that small stocks have larger average returns than big stocks, and high book-to-market (value) stocks have larger average returns than low book-to-market (growth) stocks. Similarly, portfolios of past winners (high momentum) have larger average returns than portfolios of past losers (low momentum). We have documented in the previous section that our GDP-based three-factor model prices the size/book-to-market and momentum portfolios well. Thus, our three factors must be capturing the size, value and momentum premium. To shed more light on this issue we estimate time-series regressions in which we regress the returns of the 35 size/book-to-market and momentum-sorted portfolios on the market excess return and the returns of the factor-mimicking portfolios for residential investment and business structures. The results are reported in Table 14.

The upper Panel shows the betas of the 25 size and book-to-market-sorted portfolios as well as their t-statistics. The market excess return does not generate a spread in betas that is in line with the stylized facts reported above. Small firm betas are not generally larger than large firm betas, and high book-to-market firms do not have higher betas than low book-to-market firms (they rather have lower betas, which is at odds with the value premium in observed stock returns). We thus conclude that the market excess return does not explain the size and value effect. This picture changes significantly when we consider the betas with respect to residential investment. With one exception (the lowest book-to-market quintile) the betas decrease monotonically as we move from small to large firms. Similarly, within each size quintile the betas increase almost monotonically as we move from low book-to-market portfolios to high book-to-market portfolios. We thus conclude that the residential investment-factor explains the size and the value premia. The third factor, business structures, does not explain the size and value premium. In fact, betas for large firms are larger than those for small firms, which is at odds with the size effect. There is no clear pattern with respect to the book-to-market sort.

The lower Panel of Table 14 shows the betas of the 10 momentum portfolios as well as their t-statistics. The betas with respect to the market excess returns across the ten momentum portfolios are u-shaped. Thus, the market excess return does not explain the momentum effect.

The betas with respect to residential investment are slightly more in line with the momentum effect. They tend to increase as we move from the loser portfolios to the winner portfolios. However, the spread in betas is small; the difference between the betas for the two extreme portfolios is only 0.24. The third factor, business structures, captures the momentum effect almost perfectly. The betas increase monotonically as we move from the loser portfolios to the winner portfolios, and the spread between the extreme portfolios is large (0.85).

Table 14: Factor Betas

The table shows time-series betas ($\boldsymbol{\beta}_j$) by the regression

$$R_{j,t} = a_j + \beta_{j,MKT} MKT_t + \beta_{j,RES} RES_t + \beta_{j,BST} BST_t + \epsilon_{j,t},$$

where $R_{j,t}$ are 25 Fama-French portfolios and 10 momentum portfolios (2-12). The risk factors are the market excess return (MKT), and factor mimicking portfolios for residential investment (RES) and business structures (BST). T-statistics are based on the Newey and West (1987) HAC covariance matrix with automatic lag length selection according to Andrews (1991). All data are monthly and the sample period is from January 1951 to December 2010.

25 Fama-French portfolios										
	Book-to-Market Equity					Book-to-Market Equity				
	Low	2	3	4	High	Low	2	3	4	High
	β_{MKT}					$t\,(\beta_{MKT})$				
Small	1.26	1.04	0.85	0.75	0.73	19.07	18.41	19.23	19.71	18.55
2	1.32	1.01	0.85	0.79	0.81	29.98	27.37	26.69	0.79	23.03
3	1.27	0.99	0.86	0.80	0.77	37.02	37.17	28.88	29.06	22.96
4	1.22	1.02	0.93	0.85	0.88	41.97	34.34	28.89	29.04	24.92
Big	1.01	0.97	0.93	0.84	0.85	47.44	41.42	30.05	24.33	19.34
	β_{RES}					$t\,(\beta_{RES})$				
Small	-0.07	0.13	0.27	0.38	0.54	-1.13	2.17	5.99	9.42	13.03
2	-0.19	0.09	0.24	0.37	0.53	-4.04	2.31	7.40	12.33	14.08
3	-0.17	0.08	0.20	0.32	0.48	-4.25	3.02	6.99	10.43	15.82
4	-0.17	0.04	0.15	0.25	0.36	-5.57	1.16	4.55	8.59	9.43
Big	-0.04	-0.02	-0.04	0.15	0.26	-2.01	-0.68	-1.40	4.12	6.00
	β_{BST}					$t\,(\beta_{BST})$				
Small	-0.24	-0.22	-0.24	-0.24	-0.32	-4.98	-5.65	-7.81	-9.84	-12.45
2	-0.13	-0.18	-0.18	-0.19	-0.26	-4.04	-6.40	-8.20	-9.65	-11.65
3	-0.11	-0.11	-0.12	-0.12	-0.21	-4.27	-5.43	-5.28	-5.00	-9.65
4	-0.03	-0.05	-0.07	-0.07	-0.12	-1.39	-2.55	-2.77	-3.34	-3.90
Big	-0.00	0.05	0.08	0.04	0.01	-0.25	3.31	3.72	1.74	0.31
10 momentum portfolios										
	1	2	3	4	5	1	2	3	4	5
	6	7	8	9	10	6	7	8	9	10
	β_{MKT}					$t\,(\beta_{MKT})$				
Low	1.12	0.96	0.83	0.85	0.85	27.70	33.69	29.42	29.95	31.89
High	0.92	0.91	0.99	1.07	1.32	33.86	31.95	43.79	42.46	35.11
	β_{RES}					$t\,(\beta_{RES})$				
Low	-0.18	-0.10	-0.03	0.00	0.04	-4.21	-3.70	-0.88	0.12	1.43
High	0.05	0.11	0.11	0.12	0.06	1.88	4.37	5.20	4.95	1.78
	β_{BST}					$t\,(\beta_{BST})$				
Low	-0.57	-0.40	-0.30	-0.17	-0.09	-16.82	-20.51	-9.46	-9.08	-3.92
High	-0.02	0.04	0.14	0.19	0.28	-0.82	2.38	9.24	10.92	14.08

Table 15 shows the alphas of the time-series regressions described above as well as the (unadjusted) R^2s. Of the 25 size- and book-to-market-sorted portfolios only three have a significant alpha. The three significant alphas range from 0.18% to 0.20% which corresponds to an annualized return of about 2.4%. We thus conclude that the pricing errors implied by our three-factor model are reasonably small. The R^2s tend to be larger for big stocks than for small stocks. Fifteen of the R^2s are larger than 0.80, and none is less than 0.60.

With respect to the momentum portfolios, only the winner portfolio (high momentum) has a significant (at the 5% level) alpha of 0.22%. All other alphas are insignificant and economically small (within the range of -0.20% to 0.17%). All ten R^2s are above 0.80. Notwithstanding the good performance of the GDP-based three-factor model in the individual time-series regressions, the Gibbons, Ross, and Shanken (1989) test statistic calculated across all 35 alphas rejects the model at the 1% level.

In summary, the results of the time-series regressions imply that the vanguard of the business cycle, residential investment, produces betas which are in line with the average returns of the size / book-to-market portfolios. The rear guard of the business cycle, business structures, on the other hand, produces betas in line with momentum profits. Only four of the 35 portfolios have individually significant alphas, and these alphas are rather small in economic terms.

Table 15: Factor Alphas

The table shows time-series alphas (a_j) by the regression

$$R_{j,t} = a_j + \beta_{j,MKT} MKT_t + \beta_{j,RES} RES_t + \beta_{j,BST} BST_t + \epsilon_{j,t}.$$

where $R_{j,t}$ are 25 Fama-French portfolios and 10 momentum portfolios (2-12). The risk factors are the market excess return (MKT), and factor mimicking portfolios for residential investment (RES) and business structures (BST). T-statistics are based on the Newey and West (1987) HAC covariance matrix with automatic lag length selection according to Andrews (1991). The time-series R^2 is unadjusted. GRS is the test statistic suggested by Gibbons, Ross, and Shanken (1989):

$$\frac{T-N-K}{N}[1+\hat{\boldsymbol{\mu}}'\hat{\boldsymbol{\Sigma}}_{ff}^{-1}\hat{\boldsymbol{\mu}}]^{-1}\hat{\boldsymbol{\alpha}}'\hat{\boldsymbol{\Sigma}}_{\epsilon\epsilon}^{-1}\hat{\boldsymbol{\alpha}} \sim F_{N,T-N-K}.$$

All data are monthly and the sample period is from January 1951 to December 2010.

	25 Fama-French portfolios									
	Book-to-Market Equity					Book-to-Market Equity				
	Low	2	3	4	High	Low	2	3	4	High
	a_j					$t(a_j)$				
Small	-0.18	0.16	0.09	0.18	0.13	-0.91	0.96	0.75	1.56	1.09
2	0.06	0.14	0.20	0.08	-0.02	0.42	1.36	2.17	0.88	-0.17
3	0.14	0.18	0.10	0.07	0.02	1.31	2.14	1.23	0.83	0.22
4	0.18	0.04	0.10	0.06	-0.09	1.99	0.60	1.25	0.70	-0.91
Big	0.00	0.02	0.11	-0.11	-0.18	0.04	0.25	1.36	-1.14	-1.69
	R^2									
Small	0.63	0.66	0.71	0.74	0.78					
2	0.76	0.78	0.80	0.82	0.82					
3	0.81	0.84	0.83	0.82	0.81					
4	0.87	0.87	0.84	0.82	0.77					
Big	0.88	0.88	0.80	0.75	0.67					
	10 momentum portfolios									
	1	2	3	4	5	1	2	3	4	5
	6	7	8	9	10	6	7	8	9	10
	a_j					$t(a_j)$				
Low	-0.20	0.12	0.17	0.09	-0.01	-1.66	1.23	1.83	1.17	-0.13
High	-0.01	-0.08	-0.02	-0.05	0.22	-0.17	-1.23	-0.28	-0.69	2.15
	R^2									
Low	0.86	0.89	0.85	0.85	0.85					
High	0.86	0.86	0.90	0.90	0.80					
	GRS-test: 2.21					p-value: 0.0001				

Industry Portfolios and Horse Races. As suggested by Lewellen, Nagel, and Shanken (2010), we further expand the set of test assets and include 30 industry portfolios together with

25 double sorted size/book-to-market portfolios and 10 momentum portfolios.[30] Within this demanding framework we test our three-factor model against the Carhart four-factor model. Panel A of Table 16 shows the results of four models, the CAPM, the Carhart four-factor model, a GDP-based two-factor model (market excess return and residential investment) and our GDP-based three-factor model including the market excess return, residential investment and business structures. The CAPM fails. The pricing error (the constant) is large and significant, and the estimate of the risk premium is negative. The Carhart model performs better. The risk premia for the book-to-market factor and the momentum factor are significant. However, the constant, albeit smaller than in the CAPM, is still significant.

Specification (3), the GDP-based two-factor model, shows that the market factor combined with residential investment explains the cross-section of the 65 portfolios already well. The constant is smaller than in the Carhart model (but still significant at the 10% level) and the risk premium for residential investment is positive and significant. Specification (4), our three-factor model, performs even better. The constant is 0.30% and insignificant (t: 1.17). In comparison, the constant of the Fama-French/Carhart four-factor model (specification (2)) is twice as large and significant (t: 2.79). Residential investment as well as business structures have both significant risk premia (at the 1% and 5% level, respectively). However, the estimates of the factor risk premia are not as close to their sample means than those reported in Table 13 above. The OLS R^2 is as large as 0.54 (GLS R^2: 0.26). We check for the portfolios behind the improvement (not reported) of model (4) compared to model (3). The better fit is driven by the momentum portfolios (MomL, MomH) and industry portfolios which load most (in absolute terms) on the momentum factor WML: Clothes, Textiles, Autos, Coal and Oil. Figure 7 provides a graphical impression of the pricing error implied by the GDP-based two-factor and three-factor model.

Finally, we test whether SMB, HML, and WML contain pricing relevant information beyond

[30] We use the same set of test assets as Adrian, Etula, and Muir (2011) (but our time-series is longer). Adrian, Etula, and Muir show that "shocks to the aggregate leverage of security broker-dealers" can explain the returns of the 65 test assets well. They find R^2s ranging from 0.45 (single factor) to 0.55 (adding the three Fama-French factors and the momentum factor).

that contained in our GDP-based three-factor model. We orthogonalize the Carhart factors (SMB, HML, WML) with respect to the GDP mimicking portfolios RES and BST and include the orthogonalized factors in an extended model.[31] We find that SDF loadings as well as risk premia for SMB and HML are not significant once we account for the GDP-based factors. The factor price for WML is substantially reduced in economic terms, from 0.72% in specification (4) to 0.19%, but is still significant (t: 2.05). The results for RES and BST are mainly unchanged. Therefore, it seems fair to conclude that the GDP-based three factor model captures most of the pricing relevant information inherent in the traditional four-factor model.

[31]We follow Ferguson and Shockley (2003) to extract the orthogonal portion of the SMB, HML, and WML factors. For example, we regress HML on RES and BST, and collect the time-series residuals plus the estimated intercept.

Table 16: GDP Mimicking Portfolio Factors and 65 Equity Portfolios

The table reports GMM estimates of asset pricing models given by

$$E(\mathbf{R}_t) = \gamma + \boldsymbol{\beta}\boldsymbol{\lambda}.$$

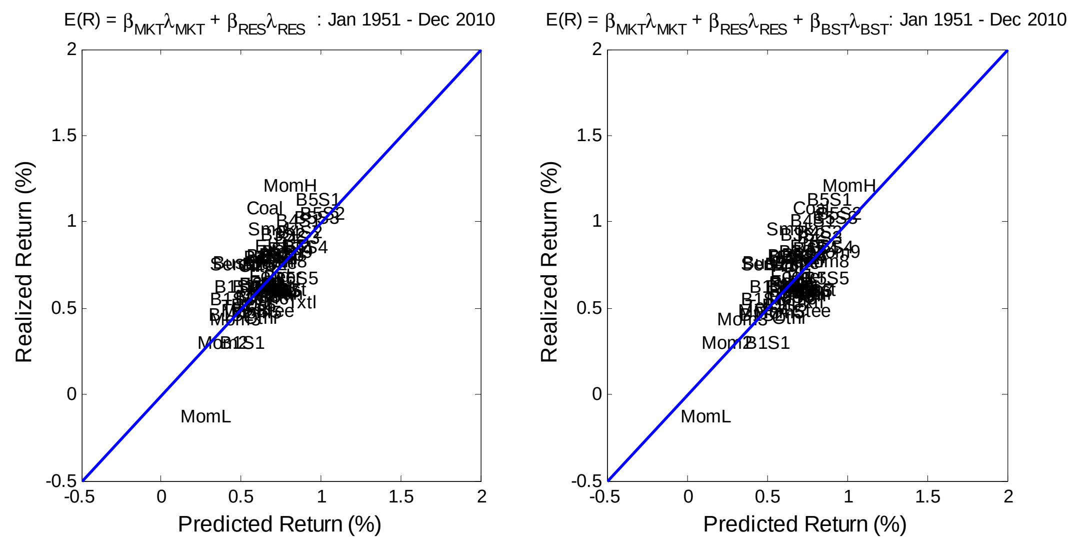
Estimates of the SDF loadings $\mathbf{b}$ and a constant γ are obtained by exploiting the moment restrictions $E(\mathbf{R}_t[1 - (\mathbf{f}_t - \boldsymbol{\mu})'\mathbf{b}] - \gamma) = 0$ and $E(\mathbf{f}_t - \boldsymbol{\mu}) = 0$, where $\mathbf{R}_t$ is a vector of stock returns and $\mathbf{f}_t$ is a vector of risk factors. The loadings on the stochastic discount factor are multiplied by 100 in the table. Factor risk prices are calculated as $\boldsymbol{\lambda} = \boldsymbol{\Sigma}_{ff}\mathbf{b}$, standard errors are obtained by the delta method, and are corrected for the fact that the covariance matrix of risk factors $\boldsymbol{\Sigma}_{ff}$ is estimated. The stock returns are 25 Fama-French portfolios sorted by size and book-to-market, 10 momentum portfolios (2-12) and 30 industry portfolios. In Panel A, the risk factors are the Fama-French / Carhart factors and the factor mimicking portfolios for residential investment (RES) and business structures (BST). In Panel B, the risk factors are RES and BST and in addition portion of the SMB, HML, WML factors orthogonal to RES and BST (as described in Ferguson and Shockley (2003)). Estimation is by first-stage GMM using the identity matrix. T-statistics are reported in parentheses and are based on the Newey and West (1987) HAC covariance matrix with automatic lag length selection according to Andrews (1991). We report the cross-sectional OLS R^2, below the cross-sectional GLS R^2, the mean absolute error (MAE), and the Hansen-Jagannathan distance (HJ) with the simulation-based p-value for a test of a zero HJ. All data are monthly and the sample period is from January 1951 to December 2010.

Panel A:		Traditional Four-Factor Model and the GDP-based Model									
		γ	MKT	SMB	HML	WML	RES	BST	$R^{2\,OLS}_{GLS}$	MAE	HJ(pv)
(1)	$\boldsymbol{b}$	0.88	-1.04						0.02	0.17	0.44
		(3.89)	(-0.65)						0.01		(0.00)
	$\boldsymbol{\lambda}$		-0.20								
			(-0.65)								
(2)	$\boldsymbol{b}$	0.68	0.71	2.38	4.97	5.18			0.59	0.11	0.40
		(2.79)	(0.38)	(1.56)	(2.58)	(3.37)			0.24		(0.00)
	$\boldsymbol{\lambda}$		-0.04	0.12	0.22	0.72					
			(-0.13)	(1.04)	(1.79)	(4.35)					
(3)	$\boldsymbol{b}$	0.46	-0.94				6.32		0.46	0.12	0.39
		(1.82)	(-0.53)				(4.31)		0.26		(0.00)
	$\boldsymbol{\lambda}$		0.15				0.96				
			(0.45)				(3.92)				
(4)	$\boldsymbol{b}$	0.30	1.18				5.81	1.57	0.54	0.12	0.39
		(1.17)	(0.60)				(3.69)	(1.49)	0.26		(0.00)
	$\boldsymbol{\lambda}$		0.35				1.04	0.69			
			(1.09)				(4.43)	(2.48)			

Panel B:		SMB, HML and WML orthogonal to RES and BST									
		γ	MKT	SMB⊥	HML⊥	WML⊥	RES	BST	$R^{2\,OLS}_{GLS}$	MAE	HJ(pv)
(5)	$\boldsymbol{b}$	0.56	0.79	-6.91	20.49	30.49	5.35	1.50	0.66	0.10	0.38
		(2.30)	(0.24)	(-1.31)	(1.43)	(1.76)	(2.67)	(1.12)	0.33		(0.01)
	$\boldsymbol{\lambda}$		0.08	0.17	-0.14	0.19	0.95	0.68			
			(0.25)	(1.57)	(-1.26)	(2.05)	(4.40)	(2.42)			

Figure 7: Cross-Section of 65 Equity Portfolios

The figure shows pricing errors for 25 Fama-French portfolios sorted by size and book-to-market, 10 momentum portfolios (2-12), and 30 industry portfolios using monthly GDP mimicking portfolios. RES is residential investment, BST is business structures, and MKT is the market excess return. The sample period is from January 1951 to December 2010.



2.6 Robustness

We have performed a large number of robustness test. In order to economize on space we do not report the results in this paper but rather present them in Appendix B.3. This section briefly describes the additional test we have performed and their main results.

In the main text we use as test assets 25 size and book-to-market sorted portfolios, 10 momentum portfolios, the combination of these two sets, and and the extended set plus 30 industry portfolios. We perform similar tests for 25 portfolios sorted on size and momentum. The results are very similar to those presented in Table 5 of the paper.

In the paper we showed the results of our GDP-based three-factor model only for the extended set of 35 test assets and the full set which also includes the industry portfolios. We have also estimated three-factor model on the 25 size and book-to-market sorted portfolios and the 10 momentum portfolios. The results are similar to those presented in the paper.

Vassalou (2003) constructs a mimicking portfolios that tracks future GDP. She then shows that this factor prices 25 portfolios sorted on size and book-to-market about as well as the

Fama/French three-factor model. Constructing a mimicking portfolio which predicts future changes (rather than contemporaneous changes) in GDP or its components is an obvious way to eliminate the lag in some of the GDP components and aggregate GDP. We therefore adopt the approach of Vassalou (2003). In line with our expectations we find that the performance of the lagging GDP components improves significantly. In particular the estimated factor risk premia are positive. These findings corroborate our conclusion that the lead-lag structure in aggregate GDP and its components has important asset pricing implications.

When we split our samples in two parts (1951-1980 and 1981-2010) we obtain results that are comparable to those for the full sample shown in the text.

We construct our mimicking portfolio over the entire data set and then test the ability of the mimicking portfolios to price the test assets. One potential objection against this procedure is that the data used to construct the mimicking portfolios was unavailable during the sample period. We therefore also perform out-of-sample tests. We use the data from 1951-1980 to construct the mimicking portfolios and then test whether these mimicking portfolios price our test assets in 1981. We then proceed using an expanding windows approach, i.e. we next use the data from 1951-1981 to construct mimicking portfolios and then test their cross-sectional pricing ability for the return of the test assets in 1982, and so on. The results of the out-of-sample tests are very similar to those that we obtain when we apply the in-sample procedure to the same sample period (i.e. 1981-2010).

The mimicking portfolios for the six GDP components are constructed independently. For each component we identify the weights of the portfolio that has maximum correlation with the growth rate of the respective GDP component. Alternatively we first estimate a first-order VAR system of the five GDP components. We then construct mimicking portfolios that have maximum correlation with the VAR innovations. When we use this alternative set of factor-mimicking portfolios we obtain results which are similar to those presented in the text.

In the paper we use annual GDP data to construct the mimicking portfolios. We repeated the analysis using quarterly data instead. The results for durable consumption are a bit odd (the mimicking portfolio has extreme weights and the results of the asset pricing tests differ significantly from those obtained using annual data). The results for the other four components are similar to those presented in the paper and yield the same conclusions.

2.7 Conclusion

Relating stock market returns to GDP has been the topic of many previous studies in financial economics. Empirical results have been mixed at best. We add to this literature by taking the lead-lag structure of GDP components into account. We find that the lead-lag structure of GDP components is mirrored in the factor structure of size, book-to-market, and momentum portfolio returns. The leading GDP components - residential investment in particular - explain the returns on size and book-to-market portfolios very well. The lagging GDP components (business structures in particular), on the other hand, capture momentum returns surprisingly well but completely fail to explain the returns on size and book-to-market sorted portfolios.

Based on these results, we propose and test a GDP-based three-factor model with the market excess return, one leading GDP component, and one lagging GDP component. This model accounts for the lead-lag structure of GDP. We find that our GDP-based three-factor model is able to explain the cross-section of returns for 65 stock portfolios sorted on size, book-to-market, momentum, and industry at least as well as the Fama-French/Carhart four-factor model.

Chapter 3

3 International Diversification Benefits with Foreign Exchange Investment Styles

This chapter is coauthored by Felix Schindler, and Andreas Schrimpf.

3.1 Introduction

Unsurprisingly, the discovery of various foreign exchange (FX) market "anomalies" (e.g. Froot and Thaler, 1990) since the end of the Bretton Woods period has generated considerable interest by researchers and investors. Similar to value and momentum in equity markets, various investment styles have been devised to exploit the market anomalies documented in FX markets. While style-based investments and their role for portfolio allocation in equity markets have been widely studied (e.g. Eun, Huang, and Lai, 2008, or Eun and Lee, 2010a), however, there is considerably less knowledge about the portfolio implications of style investing in FX markets.

In this paper, we provide a comprehensive analysis of portfolio choice with three widely practiced FX investment styles. We go beyond well-known carry trades and investigate two further popular FX investment strategies, namely FX momentum and FX value. The most prominent strategy is arguably the carry trade, which is the trading strategy derived from the "forward premium puzzle" (Fama, 1984) and consists of long positions in high interest rate currencies funded by borrowing in low interest rate currencies. If the interest differential is not offset by corresponding exchange rate movements (i.e. if uncovered interest rate parity, UIP, does not hold), there are considerable gains to be made from this form of currency speculation. Carry trades have shown to be highly profitable (see, e.g. Lustig and Verdelhan, 2007) and are widely used among professional currency fund managers (Pojarliev and Levich, 2008). The remaining two strategies we study (FX momentum, and FX value) have also shown to be profitable as documented in the recent work by Ang and Chen (2010), Asness, Moskowitz, and Pedersen (2012), Burnside, Eichenbaum, and Rebelo (2011) or Menkhoff, Sarno, Schmeling, and Schrimpf (2012b), for instance. As pointed out by Pojarliev and Levich (2008), these additional strategies are also very popular in the asset management industry.[32]

Given the Sharpe ratios of these FX strategies documented in the extant literature as well as their fairly low correlations to traditional asset classes, one straightforward question arises: Does the addition of FX styles to the investment opportunity set improve the risk-return profile of international well-diversified portfolios? There is more to this question than meets the eye. The unhedged return of an international asset is composed of two parts, the core asset return in local currency and the currency return. Taking the recent evidence on common factors in FX markets seriously (Lustig, Roussanov, and Verdelhan, 2011, Verdelhan, 2011), investors should be able to derive diversification benefits by considering both components of international investing. In fact, this is in line with modern portfolio management practice, which often relies on a two-layer approach (see Pojarliev and Levich, 2012, for a detailed description). Typically, in a first step, a hedging strategy is implemented to remove "unintended" currency exposure in core assets such as bonds or equities. In a second step, a return-seeking overlay is implemented and strategic positions in currency markets are established (e.g. Olma and Siegel, 2004).[33]

While optimal currency hedging (i.e. the first step) has received considerable attention in the academic literature so far, e.g. most recently in the work by Campbell, Sunderam, and Viceira (2010), there are hardly any academic studies about the second step of modern currency management, that is, strategic currency investing. This paper aims at providing a

[32] Just to name a few real-life examples which are available for non-institutional investors: Deutsche Bank Carry ETF, Deutsche Bank Momentum ETF, and Deutsche Bank Valuation ETF (all based on the G10 currencies); Investment products with carry strategies including emerging market currencies are also available from UBS (V24 Carry TR Index).

[33] Several leading asset managers have recently added new forms of FX investment style-related currency management schemes to their product portfolios (which are similar to our set up), stressing the economic importance and relevance of the question we analyze. See, inter alia, UBS: "Currency Management in Equity Portfolios", 2006; Deutsche Bank: "Currency Indices in a Portfolio Context", 2007; J.P. Morgan: "Managing Currency", 2008; BNY Mellon: "New Approaches to Global Currency Management", 2010; PIMCO: "Asset Allocation Rx for Fx: A Risk Factor-Based Approach to Currencies", 2011. Russel Investments: "Conscious Currency - A new approach to understanding currency exposure", 2011. We thank Stephan Siegel who pointed us to this trend.

better understanding of the most popular FX investment strategies and their joint stochastic properties with global bonds and equities, and therefore is a first step towards filling this gap. As the benchmark allocation, we consider a well-diversified portfolio of global bonds and stocks. To account for FX risk exposure, we carefully hedge our benchmark assets against exchange rate risk before testing the benefits from FX style investing. The hedging strategies for the benchmark assets – a full hedge, an optimal hedge and a conditional optimal hedge – draw on recent work by Campbell, Sunderam, and Viceira (2010). This procedure allows us to differentiate between hedging benefits and speculative benefits from FX investing.[34]

Based on this empirical setup we establish three major findings. First, we show that considerable improvements in the portfolio allocation can be achieved by style investing in FX markets. Considering all three baseline FX styles raises the (annualized) Sharpe ratio from 1.25 (benchmark assets, conditional optimal hedge) to 1.62 (FX style augmented portfolio), an increase of about 30%. When the FX investment strategies are constructed from a smaller set of developed market (G10) currencies, we still find a significant increase of the Sharpe ratio by 17%. These results hold after correcting for transaction costs (based on quoted bid-ask spreads) that are implied by re-balancing of the FX style portfolios. Furthermore, in an extensive out-of-sample evaluation, we find that the diversification benefits (measured by the increase of the Sharpe ratio) remain significant in this more realistic setting where portfolio decisions are taken in real-time.

Second, we find that considering combinations of all three FX strategies improves the portfolio allocations more than relying on a single FX style in isolation. Correlations between the FX strategies are fairly low. Thus, there are benefits from diversifying within the space of FX investment strategies (a point which is also made by Jorda and Taylor, 2012).[35] In our robustness tests, we also assess the benefits from FX style investing relative to benchmark assets that rely

[34]Our baseline results are obtained for FX style portfolios constructed from 30 liquid and frequently traded currencies. Moreover, we consider a reduced subset of G10 currencies and also account for transaction costs which typically occur due to a re-balancing of currency positions in the FX style portfolios.

[35]These results are in line with Eun and Lee (2010a) who show that equity style diversification is beneficial for global equity portfolios. Eun and Lee (2010a) find that global style investing in equity markets raises the annualized Sharpe ratio against a U.S. portfolio from 1.14 to 2.39. Notice that they do not account for transaction costs due to portfolio re-balancing in their baseline results.

on equity and commodity momentum and value strategies (Asness, Moskowitz, and Pedersen, 2012). We find that FX styles also provide diversification benefits in this setting with benchmark assets that are managed portfolios.

Third, we find that the improvements in portfolio Sharpe ratios from the FX style investments do not imply a worsening of the portfolio downside risk characteristics. This is notable since some of the FX strategies (especially the carry trade, see e.g. Gyntelberg and Remolona, 2007, Brunnermeier, Nagel, and Pedersen, 2009) are prone to occasional large losses, i.e. have negatively skewed return distributions. In our robustness tests, we consider three different additional checks of this issue. We start with an extreme scenario where the carry trade strategy is hedged against extreme downside risk using currency options (as in Jurek, 2009, Burnside, 2011). We find that the carry trade still provides some (albeit lower) improvements in the mean-variance space. Next, related to our second major finding, we consider combinations of the three FX strategies. These diversified FX strategies turn out to be much less exposed to negative skewness. Moreover, such combinations of FX styles provide even larger gains as measured by Sharpe ratios. As a final check, we rely on the stochastic dominance criterion, which is based on less restrictive assumptions compared to the traditional mean-variance framework. To that end, we use multivariate second-order and third-order stochastic dominance tests proposed by Post and Versijp (2007). The third-order test allows explicitly to account for investors who dislike negatively skewed return distributions at the portfolio level (i.e. it takes into account the skewness which remains after FX styles become part of a well-diversified benchmark portfolio). Also the stochastic dominance tests corroborate our baseline results.

Our paper proceeds as follows. We briefly discuss the related literature in Section 3.2. In Section 3.3, we provide a detailed description of the FX investment styles in a common framework. Section 3.4 describes our dataset and shows how our FX style portfolios and benchmark portfolios are constructed. Section 3.5 presents our major empirical results on the three FX investment strategies and illustrates the gains in international portfolio diversification that can be achieved by FX style investing. Section 3.6 provides robustness tests and looks at our baseline results from various further angles. Finally, we conclude in Section 3.7.

3.2 Related Literature

Ever since the work of Grubel (1968) and Solnik (1974), researchers have become aware of the potential benefits from international diversification. Somewhat surprisingly, many empirical studies were hardly able to identify statistically significant diversification benefits from investments in international developed equity markets.[36] A possible explanation is the ongoing integration of global markets and a thus increasing correlation among international assets (e.g. Eun and Lee, 2010b; or Christoffersen, Errunza, Jacobs, and Jin, 2012). By contrast, style-based stock market investing seems to provide distinguishable diversification benefits. Eun, Huang, and Lai (2008) show that benefits are significant for international small cap stocks, and Eun and Lee (2010a) report similar results for value and momentum strategies.

Most of the diversification studies mentioned above do not rely on returns that are hedged against currency risk. Relatively little attention is devoted to the role of the foreign exchange rate component, which is by construction an unavoidable element of international investments. Some studies, however, carefully consider the exchange rate component in foreign investments, in particular Glen and Jorion (1993) and most recently Campbell, Sunderam, and Viceira (2010). These studies first and foremost consider the role of individual currency positions and their role for international portfolios. Glen and Jorion (1993) as well as de Roon, Nijman, and Werker (2003) do not find (significant) diversification benefits of simple currency positions that go beyond fully hedging the currency risk exposure of stock and bond portfolios. Campbell, Sunderam, and Viceira (2010) find that single currencies are attractive to minimize the risk of global equities. Interestingly, all four studies find further increased portfolio Sharpe ratios for a hedging strategy conditional on the interest rate differential of the domestic country to the foreign country (hence,

[36] To our knowledge, there is not a single study analyzing (non-style-based) international developed stock market returns which finds significant diversification benefits for the tangency portfolio and a recent time period (Britten-Jones, 1999; Errunza, Hogan, and Hung, 1999; Eun, Huang, and Lai, 2008; Kan and Zhou, 2012a; Eun and Lee, 2010a, among others).

mimicking some kind of carry trade strategy).

Our study is motivated by these initial findings for simple conditional hedges based on carry in the older literature. It also draws on the results on the profitability of several FX investment strategies documented in more recent papers (e.g. Ang and Chen, 2010; Asness, Moskowitz, and Pedersen, 2012; Burnside, Eichenbaum, and Rebelo, 2011; Lustig, Roussanov, and Verdelhan, 2011; Menkhoff, Sarno, Schmeling, and Schrimpf, 2012b). We go beyond the existing literature by focusing on the diversification benefits due to the speculative component of FX investing as opposed to the gains stemming from the hedging component. Contrary to the older literature, we employ multi-currency investment strategies, which is common practice among market participants nowadays.

3.3 FX Investment Styles

We study the diversification benefits of three FX investment styles which can be considered the most common strategies by professional currency fund managers (see e.g. Pojarliev and Levich, 2008) and which have received the utmost attention in the recent academic literature. We will describe these investment strategies in the following; they are known as the currency carry trade, the FX momentum, and the FX value strategy.

In our empirical analysis, we use monthly observations and re-balance the style portfolios at the beginning of every month. The end-of-month payoff on a long forward position (also denoted as the "FX excess return" in the following) for currency j is measured as

$$RX_{j;t+1} = \frac{S_{j;t+1} - F_{j;t}}{S_{j;t}}, \tag{20}$$

where $S_{j;t}$ is the spot U.S. dollar (USD) price of one unit of foreign currency j at time $t = 0, ..., T$ and $F_{j;t}$ is the one period forward price. Computed this way, the FX return is an excess return since it is a zero net investment consisting of selling USD in the forward market for the foreign currency in t and buying USD at the future spot rate in $t + 1$. All three FX investment strategies generally rely on long-short positions in foreign currencies conditional on a specific signal available one period before. Following Asness, Moskowitz, and Pedersen (2012), we build currency portfolios by weighting them according to their time t specific cross-sectional rank:

$$
\begin{aligned}
w_{j;t}^{z(s)} &= c_t \left(rank\left(z\left(s\right)_{j;t}\right) - \sum_{j=1}^{J_t} rank\left(z\left(s\right)_{j;t}\right)/J_t \right), && (21) \\
RZ_{s;t+1} &= \sum_{j=1}^{J_t} w_{j;t}^{z(s)} RX_{j;t+1} \qquad s = \{C,\, M,\, V\}, && (22)
\end{aligned}
$$

where $RZ_{s;t+1}$ denotes the return on the style-based trading strategy and depends on the conditioning variable $z\left(s\right)_t$. J_t denotes the set of currencies that are available for investment in period t. The choice of the signal s determines the particular strategy, e.g. the carry trade (C), FX momentum (M), or FX value (V). The constant c_t is chosen such that the style portfolios are one USD long and one USD short. Using the rank instead of the particular signal mitigates outliers and reduces transaction costs due to portfolio re-balancing (see below). Our rank-based portfolios result in an entirely data-driven weighting scheme for the individual currencies. In contrast, an equally weighted approach as used in much of the extant literature (e.g. Lustig, Roussanov, and Verdelhan, 2011; Menkhoff, Sarno, Schmeling, and Schrimpf, 2012a) would give us an additional degree of freedom by choosing a particular quantile used as a cutoff to determine long-short positions.

Carry trade strategy. The carry trade exploits the well-established empirical failure of uncovered interest rate parity (UIP) known as the "forward premium puzzle" (Fama, 1984). Following the paper by Lustig and Verdelhan (2007), our carry trade strategy goes long (forward) in an equally weighted portfolio of currencies with the largest nominal short-term interest rates (investment currencies), and short (forward) in an equally weighted portfolio of currencies with the smallest nominal short-term interest rates (funding currencies). Thus, our conditioning variable in the carry trade is the interest rate differential between the foreign and the U.S. money market, which we infer from the FX forward premium/discount

$$z(C)_{j;t} = \frac{F_{j;t}}{S_{j;t}} - 1. \tag{23}$$

Carry trade strategies are very profitable, typically have quite attractive risk-return characteristics, are widely used by practitioners and there is evidence that they leave their traces in FX turnover patterns (see Galati, Heath, and McGuire, 2007). Of particular interest in the recent academic literature is whether returns on carry strategies can be explained by a risk premium or whether they should be attributed to the presence of frictions or behavioral biases.[37] In distinction to this literature, we take their returns as given and analyze if there is a (significant) demand for carry trade investments in an internationally diversified portfolio, or in other words, if an investors can improve their investment opportunity set by an investment in a carry trade strategy.

FX momentum strategy. In fact, the carry trade is not the only FX investment style discussed in the academic literature and used by professional currency fund managers. Similar to the well-known momentum returns in stock markets (e.g. Jegadeesh and Titman, 1993), momentum profits have also been documented in FX markets (see e.g. Okunev and White, 2003; Menkhoff, Sarno, Schmeling, and Schrimpf, 2012b). Menkhoff and Taylor (2007) and Pojarliev and Levich (2008) also report some evidence for the high popularity of trend-following FX strategies by professional currency fund managers. Our momentum portfolio goes long in a portfolio of currencies with the highest past cumulative returns (so-called "winners") and short in a portfolio of currencies with the lowest past returns (so-called "losers"). The conditioning variable of the momentum strategy is the cumulative return over the past three months

[37] See, for instance, Lustig and Verdelhan (2007), Bacchetta and van Wincoop (2010), Christiansen, Ranaldo, and Söderlind (2010), Verdelhan (2010), Burnside (2011), Burnside, Han, Hirshleifer, and Wang (2011), Lustig, Roussanov, and Verdelhan (2011), Menkhoff, Sarno, Schmeling, and Schrimpf (2012a), for recent contributions.

$$z(M)_{j;t} = \prod_{\tau=0}^{2} (1 + RX_{j;t-\tau}) - 1. \tag{24}$$

Menkhoff, Sarno, Schmeling, and Schrimpf (2012b) show that the momentum signal in currency markets is stronger for more recent past returns. However, they also report that since one-month past returns are more volatile than longer-horizon past returns, transactions costs due to portfolio re-balancing are larger for the former. To take this trade-off into account, we rely on three-month past returns for our baseline FX momentum strategy. Further variations of the FX momentum strategy are covered in Appendix C.1.

FX value strategy. The basic idea behind the value strategy is to buy currencies considered to trade below a fundamental value and to sell currencies which trade above a fundamental value. One may interpret this strategy as a contrarian or long-term reversal strategy. A widely used measure for fundamental value in currency markets is the real exchange rate defined as

$$Q_{j;t} = \frac{S_{j;t} P_{j;t}}{P_t^*}, \tag{25}$$

where $P_{j;t}$ is the price level of consumer goods in country j in the local currency, and P_t^* the corresponding U.S. price level in USD. If purchasing power parity (PPP) holds between two countries, Equation (25) should be equal to one. Hence, currencies with real exchange rates below (above) unity may be regarded as "undervalued" ("overvalued"). PPP is a rather strong assumption, as an equilibrium real exchange rate can easily deviate from unity (Harrod-Balassa-Samuelson effects). Thus, to avoid the problem of defining an equilibrium real exchange rate, we use a measure of "value" defined as minus one times the cumulative five-year change of the real exchange rate as our conditioning variable

$$z(V)_{j;t} = \left(\frac{Q_{j;t-3}}{\bar{Q}_{j;t-60}} - 1 \right) \times (-1), \tag{26}$$

where $\bar{Q}_{j;t-60}$ is the real exchange rate measured as the average over a period between 5.5 and

4.5 years back in the past.[38] To avoid overlap between the momentum and value conditioning variables, we skip changes of the real exchange rate of the past three months when constructing the FX value measure. Further details on the construction of the value strategy are provided in in Appendix C.1.

3.4 Data and Portfolio Construction

FX data. Bid and ask quotes for spot and one-month forward exchange rates against the USD are from Barclays Bank International (BBI) and WM/Reuters (WMR) and are available via Thomson Reuters Datastream. The FX sample covers the 30 currencies reflecting the lion's share of global market turnover.[39] We also perform tests based on a reduced set of developed market currencies, or "G10 currencies" (currencies of Australia, Canada, Germany, Japan, New Zealand, Norway, Sweden, Switzerland, and the U.K. against the USD). We use CPI data from the IMF's International Financial Statistics (IFS) to calculate real exchange rates for the value strategy.[40] The available currency data span the period from 02/1976 to 12/2011. For the value strategy, we need 5-year changes of the real exchange rate. Thus, the sample period for the FX style returns ranges from 02/1981 to 12/2011. Appendix C.1 provides further details on the data and on some conservative data screens which we apply to ensure reliability of our data. Furthermore, we provide a detailed comparison of our FX investment style returns with similar portfolios studied by Burnside (2011), Lustig, Roussanov, and Verdelhan (2011), and Asness, Moskowitz, and Pedersen (2012).

[38]The construction of our FX value strategy is motivated by the measure studied by Asness, Moskowitz, and Pedersen (2012). The correlation between our FX value portfolio return and theirs is as large as 0.86.

[39]This sample of currencies includes Australia, Brazil, Canada, Denmark, Euro, France, Germany, Hungary, Iceland, India, Indonesia, Israel, Italy, Japan, Mexico, New Zealand, Norway, Philippines, Poland, Russia, South Africa, South Korea, Spain, Sweden, Switzerland, Taiwan, Thailand, Turkey, Ukraine, and the United Kingdom. According to BIS (2010), our set of currencies covers more than 95 percent of the global FX market turnover in April 2010.

[40]Since the CPI has an arbitrary base year unrelated to PPP, we use the PPP estimate of Heston, Summers, and Aten (2009) for the year 2000 to determine the level of the real exchange rate. The resulting conditioning variable for the value strategy is robust, as we use changes in the real exchange rate.

Transaction costs. Our style-based investment strategies involve a re-allocation of the positions in the individual currencies in every month according to the signal by the corresponding conditioning variable. Since the monthly re-balancing of the portfolios involves transaction costs, we compute returns both with and without adjusting for bid-ask spreads. Our adjustment procedure is a conservative approach of accounting for transaction costs. The bid-ask spreads in the WMR/BBI database are based on indicative quotes and are thus likely to overstate the true transaction costs of an investor (Lyons, 2001). Moreover, in practice transaction costs may be substantially lower when currency positions are rolled via FX swaps as shown by Gilmore and Hayashi (2011).[41] Details on the transaction cost adjustment are given in Appendix C.1.

Benchmark assets: Global bonds and stocks. To quantify diversification benefits from style-based FX investing we consider a typical well-diversified international portfolio allocation consisting of global bonds and stocks. The selected countries are the same as those covered by the G10 currencies. In particular, we take the MSCI Total Return (Standard Country) stock market indices for Australia, Canada, Germany, Japan, New Zealand, Norway, Sweden, Switzerland, the United Kingdom, and the United States. These returns are measured in USD and are available for the period from 02/1981 to 12/2011. We obtain these data from Thomsen Reuters Datastream. There are no official bond market indices covering the time period and the cross-section of countries required for our study. Following Campbell, Lo, and MacKinlay (1997) and Campbell, Sunderam, and Viceira (2010), we use a log yield-return approximation to derive bond market returns. Stock and bond market returns are monthly simple returns in excess of the one-month U.S. Treasury bill rate from Ibbotson (available on the website of Kenneth R. French).[42] Our global bond and equity benchmark assets provide the best possible comparison

[41] The literature on international diversification has largely ignored the effects of transaction costs, most likely since no adequate data are available for international stock markets, which are the subject of most studies. An exception, for instance, is the work of de Roon and Nijman (2001) who incorporate transaction costs in their study of emerging market equities. They adjust for hypothetical transaction costs and show that for most countries even a small amount of transaction costs is sufficient to keep investors out of the market.

[42] In Section 3.6, we also provide detailed results with alternative benchmark assets covering momentum and value strategies in stocks and commodities (Asness, Moskowitz, and Pedersen, 2012).

to the existing literature on international diversification in developed markets.[43]

Dissecting FX speculation from FX hedging. In our analysis we carefully account for the exchange rate risk exposure inherent in the benchmark assets in order to dissect the benefits of hedging and strategic speculative currency positions. The international bond and stock market returns in our benchmark portfolio are exposed to exchange rate risk, i.e. these are unhedged returns $R^{uh}_{j;t}$. However, it is easy to counteract the exchange rate risk using a currency hedging strategy. Indeed, the effects of various hedging strategies are covered by a fairly large literature (e.g., Anderson and Danthine, 1981; Glen and Jorion, 1993, Jorion, 1994 and Campbell, Sunderam, and Viceira, 2010). Thus, it is interesting to see how various currency hedging strategies affect the benefits from "speculative" FX style investing. The currency hedged return $R^{h}_{j;t}$ can be written as

$$R^{h}_{j;t} = R^{uh}_{j;t} + \widetilde{\mathbf{\Psi}}'_{RM}\mathbf{H}_t, \tag{27}$$

where $\widetilde{\mathbf{\Psi}}'_{RM}$ is a vector of hedging positions, or risk management demands (Campbell, Sunderam, and Viceira, 2010), and $\mathbf{H}_t$ is a vector with the returns of the hedges. For the U.S. asset (bond or stock), we define $j = 0$ and for the foreign assets $j = 1, ..., 9$. In our baseline specification, the hedges are the G10 currency excess returns ($\mathbf{H}_t = \mathbf{RX}_t$). The vector of G10 currencies follows the same order as the foreign assets, i.e. $\mathbf{RX}_t = [RX_{1,t}, ..., RX_{9,t}]'$.[44]

We consider three types of hedging strategies (full hedge, optimal hedge, and conditionally optimal hedge). A simple (but popular) strategy is the *full hedge*. In this case, the investor hedges the full amount invested in a foreign asset with a counter position in the currency to which the investment is exposed. Accordingly, the vector of risk management demands for the jth foreign asset is simply $\widetilde{\mathbf{\Psi}}_{RM} = -\mathbf{1}_j$, which denotes a position in currency j of -1 and zero for

[43] Glen and Jorion (1993), de Santis and Gerard (1997), Britten-Jones (1999), de Roon, Nijman, and Werker (2003), Kan and Zhou (2012a), Eun and Lee (2010a), among others, cover similar assets and international markets.

[44] For example, if $j = 1$ corresponds to the Australian bond (or stock) market return $R^{uh}_{1;t}$, then, $RX_{1,t}$ is the currency excess return of the Australian dollar against the USD.

all other currencies. Campbell, Sunderam, and Viceira (2010) study *optimal currency hedges*. They calculate risk management demands such that the variance of U.S. and foreign bonds and stocks is minimized. Here, the vector of risk management demands is minus one times the beta coefficients of a regression of all j currency excess returns ($\mathbf{RX}_t$) on the fully hedged asset (stocks or bonds), which we denote $\widetilde{\mathbf{\Psi}}_{RM} = -\mathbf{B}_{\mathbf{RX}}$. Furthermore, Campbell, Sunderam, and Viceira (2010) propose a *conditional optimal hedge*. This strategy allows for time-varying hedging positions conditional on a specific signal. Campbell, Sunderam, and Viceira (2010) use foreign interest rate spreads, i.e. this strategy also hedges against carry trade risk. To mimick a conditional optimal hedge, we include the FX investment styles as hedges, i.e. $\mathbf{H}_t = [\mathbf{RX}_t', \mathbf{RZ}_t']'$ where the vector $\mathbf{RZ}_t$ contains the FX investment styles we want to test, and accordingly we denote the risk management demands as $\widetilde{\mathbf{\Psi}}_{RM} = [-\mathbf{B}_{\mathbf{RX}}, -\mathbf{B}_{\mathbf{RZ}}]'$.[45]

We can exploit the results based on the optimally hedged benchmark assets to dissect the improvement of the slope of the mean-variance frontier into a speculative component and a hedging component. The intuition is simple. The hedging component is the part of the diversification benefits driven by the non-zero betas (correlations) between the test assets and the benchmark assets. The speculation component is the residual, and will be driven by the return and variance of the test assets. The optimal hedging positions simply orthogonalize the benchmark assets with respect to the hedges by construction. Thus, when the benchmark assets are orthogonalized with respect to the test assets we are interested in, as is the case for the conditional optimal hedge, any improvement of the Sharpe ratio must be driven solely by the speculative component. In other words, the results we report for optimal conditional hedged benchmark assets allow to make a clear cut between the currency hedging benefits documented by Campbell, Sunderam, and Viceira (2010) and the benefits from speculative style investing we are interested in. Ap-

[45] Campbell, Sunderam, and Viceira (2010) propose two ways to implement a conditional currency hedge. First, in the main paper, they allow the single currency risk management demands ($-\mathbf{B}_{\mathbf{RX}}$) to be time-varying conditional on foreign interest rate spreads. Second, in the online appendix, they include the carry trade as a hedge together with the single currencies and consider constant risk management demands for the expanded set of hedges. Of course, the carry trade portfolio can be thought of as a managed portfolio exploiting information in foreign interest rate spreads (Cochrane, 2005, Chapter 8). Campbell, Sunderam, and Viceira (2010) find that the second approach provides very similar - often even slightly better - hedging results.

pendix C.2 covers further analytical details on the distinction between the speculation and the hedging component.

3.5 Empirical Results

3.5.1 Risk and Return Characteristics

Table 28 reports the annualized risk and return characteristics of our benchmark assets.[46] Global stocks deliver larger annualized average excess returns than bonds, 6.3% compared to 4.8% (unhedged). However, in our sample, the global bonds generate higher reward-per-risk measured by the Sharpe ratio, reflecting an extraordinarily good performance of bond markets over the past decades compared to the longer history.[47] The right-hand side of the table reports characteristics for fully hedged global assets. We find considerably smaller standard deviations for hedged global bonds and stocks compared to their unhedged counterparts.

Appendix C.4 also provides results on the effects of the optimal and the conditionally optimal hedges. In line with Campbell, Sunderam, and Viceira (2010), we find that the full hedge is risk minimizing for global bonds. For global stocks, there are significant additional benefits from an optimal hedge (using single currency positions). Finally, the conditional optimal hedge (including hedging positions in FX styles) generally does not lead to further significant improvements in the allocation for global bonds and stocks.

[46] In most of the empirical part, we consider individual global bond and stock market returns. To conserve space, the descriptive statistics in Table 28 are based on portfolios of global bonds and global stocks (GDP at PPP weights) rather than of individual country returns. The characteristics of the individual country returns are reported in Appendix C.1.

[47] The macroeconomic moderation and disinflation starting in the mid-1980s and a reduction in required risk premiums likely contributed to the extraordinary bond market performance (see Palazzo and Nobili (2010), for a discussion).

Table 17: Risk and Return Characteristics of FX Investment Styles

The table reports the mean (in percentage points), standard deviation (Std), skewness (Skew), first order autocorrelation (Ac1), and the Sharpe ratio (SR) of benchmark assets and FX investment styles. The mean is annualized by multiplying by 12, the standard deviation and the Sharpe ratio (SR) are annualized by multiplying by $\sqrt{12}$. FX styles are conditional on time t-1 forward discounts (Carry trade, "C"), the previous 3-month cumulative returns (FX momentum, "M"), and the 5-year change of the real exchange rate (FX value, "V"). The FX investment styles are based on a broad set of up to 30 currencies ("all currencies") or a smaller sub-set of major currencies ("G10 currencies"). The G10 currencies cover the same countries as the bond and stock markets. Results for FX styles are reported before and after taking into account transaction costs. The "ba/loss" gives the reduction in mean returns due to the bid-ask spread adjustment (in %). Statistics for global bonds ("B") and global stocks ("S") are based on a GDP weighted portfolio of ten international markets covering Australia, Canada, Germany, Japan, New Zealand, Norway, Sweden, Switzerland, the U.K, and the U.S. The GDP weights correspond to the countries' share in world GDP at PPP in 1980 (OECD database). All returns are measured as excess returns. Correlations are reported for returns after transaction costs. The sample period is 02/1981 - 12/2011 (371 observations).

	Mean	Std	Skew	Ac1	SR	Mean	Std	Skew	Ac1	SR	ba/loss
	no hedge					full hedge					
U.S. bonds	4.56	7.31	0.50	0.28	0.62						
U.S. stocks	6.56	15.53	-0.60	0.06	0.42						
Global bonds	4.75	7.59	0.25	0.24	0.63	4.21	5.38	0.42	0.34	0.78	
Global stocks	6.34	15.12	-0.75	0.11	0.42	5.81	14.31	-1.03	0.12	0.41	
	FX styles based on all currencies										
	before transaction costs					after transaction costs					
Carry trade	6.18	7.50	-0.63	0.10	0.82	5.67	7.52	-0.63	0.10	0.75	-8.3
FX momentum	5.34	7.68	0.25	-0.09	0.70	3.80	7.76	0.24	-0.08	0.49	-28.9
FX value	4.18	6.69	-0.31	0.09	0.62	3.82	6.69	-0.32	0.09	0.57	-8.6
	FX styles based on G10 currencies										
	before transaction costs					after transaction costs					
Carry trade	5.05	8.76	-0.70	0.03	0.58	4.71	8.78	-0.69	0.03	0.54	-6.88
FX momentum	3.18	8.41	0.13	-0.11	0.38	1.90	8.49	0.12	-0.10	0.22	-40.16
FX value	4.51	8.27	-0.00	0.07	0.54	4.19	8.27	-0.00	0.07	0.51	-7.00

	Correlation (all currencies)					Correlation (G10 currencies)				
	Global		FX styles			Global		FX styles		
	B	S	C	M	V	B	S	C	M	V
Global bonds (B)	1.00					1.00				
Global stocks (S)	0.06	1.00				0.06	1.00			
Carry trade (C)	-0.20	0.27	1.00			-0.19	0.31	1.00		
FX momentum (M)	-0.01	0.00	-0.01	1.00		0.02	-0.01	0.12	1.00	
FX value (V)	-0.02	0.01	0.05	0.09	1.00	-0.02	-0.02	0.05	0.14	1.00

Below the statistics for the benchmark assets we report the risk and return characteristics for the three FX investment style portfolios: The carry trade, FX momentum, and FX value. These currency strategies are either based on up to 30 currencies ("all currencies") or a sub-set of developed market currencies ("G10 currencies"). Based on all currencies and before taking into account transaction costs, they provide annualized average returns of 6.2%, 5.3%, and 4.2% respectively for the carry trade, FX momentum, and FX value. Our transaction cost adjustment takes into account portfolio re-balancing using bid and ask quotes. Thus, a more volatile signal will lead to more transactions and will finally result in higher transaction costs. We find that the return of the carry trade after taking into account transaction costs is 5.7%, which represents a loss in the average return of about 8%. Transaction costs for the FX value strategy are of similar magnitude, leading to a transaction cost adjusted average return of 3.8%. The FX momentum signal is considerably more volatile. Transaction costs eat up as much as 29%, reducing the adjusted return to 3.8%.

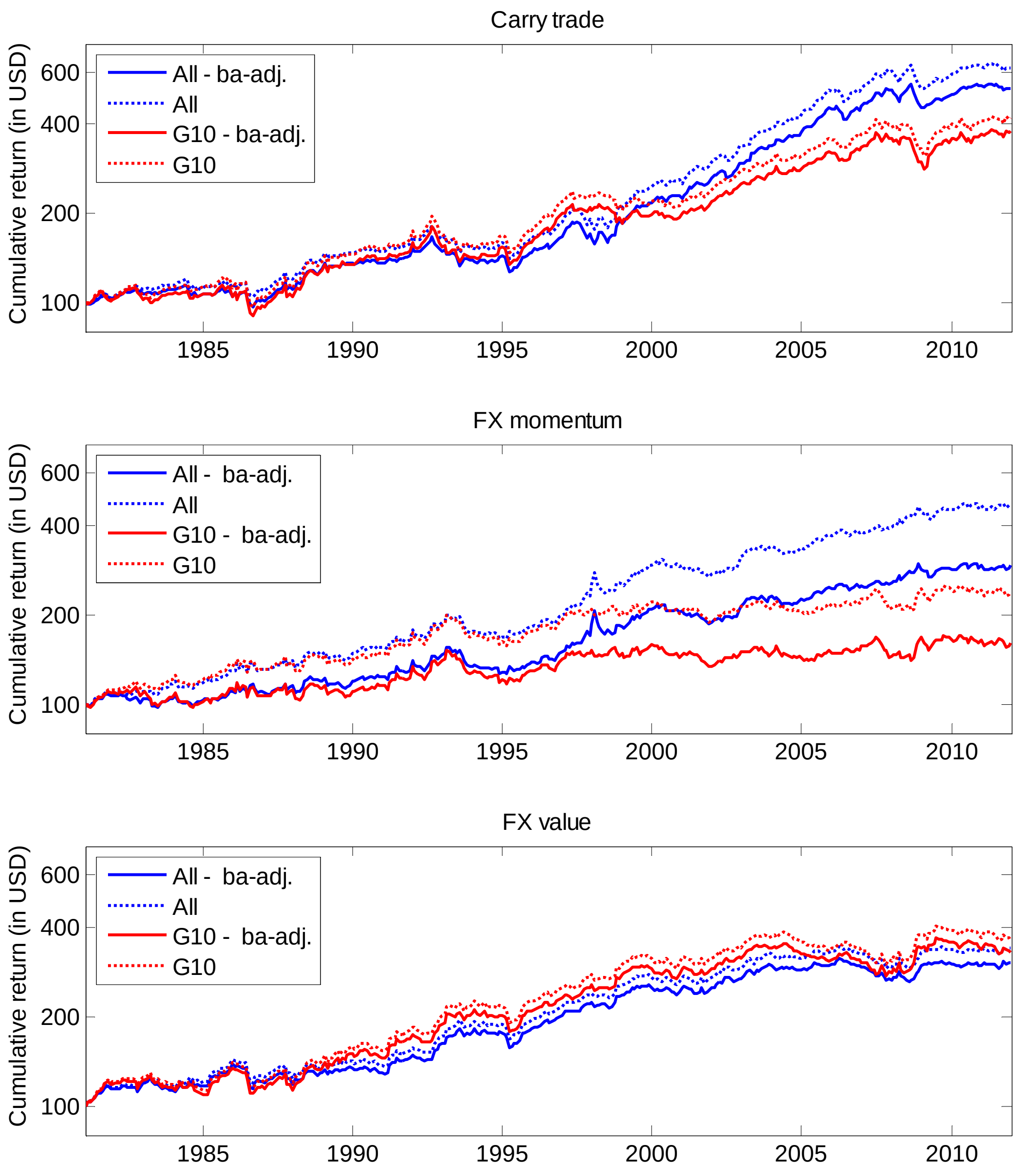
Restricting the FX investment universe to the G10 currencies leads to less diversified style portfolios (by construction) and hence larger standard deviations. In terms of average returns, the carry trade and FX value portfolios constructed from the G10 currencies are fairly similar in magnitude. The investment restriction, however, has a large negative effect on the average returns of FX momentum.[48] In addition, transaction costs are now as large as 40% and imply average returns after transaction costs of merely 1.9%. Figure 8 illustrates how the cumulative returns of the three FX styles evolve over time and how the returns are affected by transaction costs and restrictions of the investment universe to the G10 currencies.

In terms of Sharpe ratios, the FX investment styles seem to be attractive compared to global bonds and global stocks, even after accounting for transaction costs. An important observation in the table is that the return correlations of the FX portfolios and global bonds and stocks are not only fairly low, but the FX style portfolios are also barely correlated among themselves. Thus,

[48] This finding is in line with Menkhoff, Sarno, Schmeling, and Schrimpf (2012b) who show that momentum returns based on advanced economy currencies have considerably fallen over time and are no longer significant after transaction costs in recent data.

Figure 8: Cumulative Returns of FX Styles

The figure shows cumulative returns for three FX styles: The carry trade, FX momentum, and FX value. The returns are reported before and after adjusting for transaction costs due to portfolio re-balancing using bid-ask spreads ("ba-adj."). FX styles are based on up to 30 currencies or the G10 currencies. The sample period is 02/1981 - 12/2011.



strategy diversification should be beneficial in terms of mean and variance. Furthermore, it is well known that the carry trade is exposed to negative skewness, similar in magnitude to global stock markets (e.g. Gyntelberg and Remolona (2007), Brunnermeier, Nagel, and Pedersen (2009)). By contrast, returns to the FX momentum strategy exhibit a small positive skewness, and the skewness of the FX value strategy is considerably smaller in absolute terms. It therefore seems reasonable that combinations of FX style investments may also improve portfolio characteristics in terms of higher moments. We will elaborate on this point later in more depth (Section 3.6).

3.5.2 Mean-Variance Efficiency Tests

Methods. Our major interest is whether adding FX styles to an internationally diversified portfolio improves the mean-variance frontier and may thus be beneficial from an investor's perspective. With the ability to borrow and invest in a risk-free asset, a test of mean-variance efficiency comes down to a test of a shift of the tangency portfolio, or in other words, to testing if the two mean-variance frontiers intersect at the point with the maximum Sharpe ratio. Following Glen and Jorion (1993), Harvey (1995), Eun, Huang, and Lai (2008), and Eun and Lee (2010a) among others, we use frontier intersection tests to analyze if FX investment styles *significantly* shift the investment opportunity set of a global bond and equity portfolio. We estimate the following pricing model:

$$RZ_{n;t} = \alpha_n + \mathbf{R}_t'\boldsymbol{\beta}_n + \varepsilon_{n;t}, \quad t = 1, ..., T, \tag{28}$$

where $RZ_{n;t}$ is the return of $n = 1, ..., N$ test asset excess returns (FX styles), and $\mathbf{R}_t$ is a $K \times 1$ vector of benchmark asset excess returns (global bonds and global stocks). The null hypothesis of intersection is equivalent to the hypothesis that the N intercepts α_n are not significantly different from zero ($H_0 : \boldsymbol{\alpha} = \mathbf{0}_N$).[49] We report an asymptotic Wald test, $W \sim \chi^2_N$, which is robust against heteroscedasticity and autocorrelation (HAC). We use the Newey and West (1987) kernel and the automatic lag length selection procedure proposed by Andrews

[49] See Huberman and Kandel (1987), Gibbons, Ross, and Shanken (1989) and Jobson and Korkie (1989).

(1991).

Bekaert and Urias (1996) propose an alternative to the regression-based mean-variance efficiency tests, which exploits the duality between Hansen and Jagannathan (1991) bounds and mean-variance frontiers. We consider the general asset pricing restriction for the $N+K$ asset excess returns, $\widetilde{\mathbf{R}}_t = [\mathbf{RZ}_t', \mathbf{R}_t']'$:

$$E\left(\widetilde{\mathbf{R}}_t m_t\right) = \mathbf{0}_{N+K}, \tag{29}$$

where m_t is the projection of a stochastic discount factor (SDF) with mean $v = E(m_t)$ onto the demeaned $N+K$ asset returns

$$m_t = v + \left[\widetilde{\mathbf{R}}_t - E\left(\widetilde{\mathbf{R}}_t\right)\right]'\mathbf{b}. \tag{30}$$

The SDF given by Equation (30) prices the $N+K$ asset returns correctly by construction. We can write the vector of SDF coefficients as $\mathbf{b} = [\mathbf{b}_N, \mathbf{b}_K]'$. Bekaert and Urias (1996) show that mean-variance efficiency of the K benchmark assets is implied by the N restrictions $\mathbf{b}_N = \mathbf{0}_N$. Put differently, only the benchmark assets are necessary to price the augmented set of $N+K$ assets. This estimation problem can be cast in a typical Generalized Methods of Moments (GMM) framework. We set $v = 1$, which corresponds to testing intersection at the tangency portfolio in the mean-variance space, and report an HAC-robust asymptotic GMM Wald test, $SDF \sim \chi^2_N$.[50]

Results. We consider whether style-based FX investments provide diversification gains relative to ten global bond markets and ten global stock markets serving as the benchmark (K=20). We test each of the three FX styles, namely, the carry trade, FX momentum, and FX value, separately against the benchmark assets (N=1) and we test all three FX styles together against

[50]Kan and Zhou (2012a) present some evidence on the power and size of the SDF-Wald test and perform a comparison to the regression-based approach. They find no important differences between the asymptotic test statistics when returns follow a multivariate normal distribution. However, their simulation study shows that the regression-based version is favorable to the SDF-based test when returns follow a multivariate Student-t distribution.

the benchmark assets (N=3). Before testing for mean-variance efficiency, we apply three different hedging schemes to thoroughly account for the exchange risk exposure inherent in the benchmark assets. The hedging strategies follow previous work by Campbell, Sunderam, and Viceira (2010).[51] First, we implement a simple full hedge to the benchmark. Second, we add a risk minimizing optimal hedge using single G10 currency positions. And finally, we apply a risk minimizing conditional optimal hedge using single G10 currency positions as well as hedging positions in the FX style for which we test mean-variance efficiency.

Overall, we find economically large and statistically significant diversification benefits for all three FX investment styles. In Panel A of Table 18, for the carry trade based on all available currencies, we find a highly significant increase of the annualized Sharpe ratio from 1.17 to 1.47 against the fully hedged benchmark. For the FX momentum and the FX value strategies, the increase of the Sharpe ratio is from 1.17 to 1.27. Finally, adding all three FX styles to the investment universe increases the Sharpe ratio further to 1.61, i.e. there are considerable gains in terms of the maximum achievable Sharpe ratio. From Table 18, we can also infer the effects from the optimal hedge and the conditional optimal hedge of the benchmark assets on the diversification benefits from FX style investing. The optimal hedge of the benchmark assets minimizes the risk of each individual bond and stock market investment by hedge positions in the G10 currencies. The maximum attainable Sharpe ratio of the optimal hedged benchmark assets is 1.20. We find that adding all three FX styles to the optimal hedged benchmark assets, the Sharpe ratio rises further to 1.61, which is a statistically significant increase according to the results from the intersection tests.

[51] We find that the effects of the different hedging schemes in our sample are quite similar to the results reported by Campbell, Sunderam, and Viceira (2010). Thus, we report details on our estimated risk management demands ($\widetilde{\boldsymbol{\Psi}}_{RM}$) and further tests of the economic significance of the hedging strategies in Appendix C.4.

Table 18: Mean-Variance Efficiency Tests for FX Styles

This table reports the results of mean-variance frontier intersection tests for the tangency portfolio of "K" traditional assets (benchmark) when "N" FX investment styles are added to the investment universe as test assets. The K benchmark assets are ten individual bond and stock markets from Australia, Canada, Germany, Japan, New Zealand, Norway, Sweden, Switzerland, U.K., and the U.S. (K=20). Before testing for mean-variance efficiency, we apply three different hedging schemes to the benchmark assets:

$$R^{h}_{j;t} = R^{uh}_{j;t} + \widetilde{\mathbf{\Psi}}'_{RM}\mathbf{H}_t,$$

where $R^{uh}_{j;t}$ denotes the unhedged benchmark asset return of a bond or stock market of country j and $\mathbf{H}_t$ are the hedges. The full hedge unitarily hedges the country j specific currency risk. As in Campbell, Sunderam, and Viceira (2010), the optimal hedge minimizes the standard deviation of the benchmark asset returns using all G10 currencies ($\mathbf{H}_{t+1} = \mathbf{RX}_{t+1}$). The conditional optimal hedge includes positions in the single G10 currencies and the FX investment style(s) to be tested ($\mathbf{H}_{t+1} = \left[\mathbf{RX}'_{t+1}, \mathbf{RZ}'_{t+1}\right]'$) . Each FX investment style is tested separately (N=1) and all three are tested jointly (N=3) against the benchmark assets. The table reports the p-value of a regression-based test (in column W) and the p-value of a stochastic discount factor based test (column SDF) for mean-variance efficiency. The test statistics are robust against heteroscedasticity and autocorrelation (Newey and West, 1987 kernel and automatic lag length selection according to Andrews, 1991). The Sharpe ratio is reported for the tangency portfolio of all K benchmark assets and when the N test assets are added. These values are annualized by multiplying with $\sqrt{12}$. The FX investment styles are based on a broad set of up to 30 currencies (Panel A) or a smaller sub-set of G10 currencies (Panel B) and are adjusted for transaction costs due to portfolio re-balancing. The sample period is 02/1981 - 12/2011.

Benchmark: Global bonds and global stocks										
	full hedge $\widetilde{\mathbf{\Psi}}_{RM} = -\mathbf{1_j}$			optimal hedge $\widetilde{\mathbf{\Psi}}_{RM} = -\mathbf{B_{RX}}$			conditional optimal $\widetilde{\mathbf{\Psi}}_{RM} = [-\mathbf{B_{RX}}, -\mathbf{B_{RZ}}]'$			
	W	SDF	Sharpe	W	SDF	Sharpe	W	SDF	Sharpe	
	p-value		ratio	p-value		ratio	p-value		ratio	
Panel A: FX styles based on all currencies										
		Bench.:	1.17		Bench.:	1.20			Bench.	+FXS
Carry trade	0.000	0.000	1.47	0.000	0.001	1.46	0.000	0.002	1.27	1.48
FX momentum	0.002	0.004	1.27	0.002	0.003	1.31	0.003	0.005	1.21	1.31
FX value	0.009	0.014	1.27	0.009	0.017	1.31	0.003	0.008	1.17	1.30
All	0.000	0.000	1.61	0.000	0.000	1.61	0.000	0.000	1.25	1.62
Panel B: FX styles based on G10 currencies										
		Bench.:	1.17		Bench.:	1.20			Bench.	+FXS
Carry trade	0.000	0.002	1.34	0.002	0.011	1.33	0.003	0.015	1.24	1.35
FX momentum	0.192	0.182	1.19	0.191	0.168	1.22	0.199	0.178	1.20	1.22
FX value	0.018	0.021	1.26	0.021	0.026	1.29	0.012	0.016	1.18	1.28
All	0.001	0.009	1.42	0.004	0.034	1.40	0.003	0.036	1.21	1.41

The conditional optimal hedge also incorporates a risk minimizing position in the particular

FX style for which we test mean-variance efficiency. Hence, the Sharpe ratio of the benchmark depends on the specific test asset in this case. Investors using the conditional optimal hedge increase the Sharpe ratio of the benchmark portfolio if they hedge against the carry risk (1.27), FX momentum risk (1.21), and the risk of all three FX styles (1.25). The Sharpe ratio of the benchmark portfolio hedged against FX value risk, however, is slightly decreased (1.17) compared to the simple optimal hedge (1.20). Now, what happens to the allocation if speculative FX style positions are added to the portfolio of the conditionally hedged benchmark assets? Since the benchmark is hedged against FX style risk in these specifications as discussed above, the reported diversification benefits reflect the pure speculative component of FX style investing.[52] The corresponding Sharpe ratios are reported in the right column of Table 2. Adding the FX styles to the asset allocation results in marked improvements in the maximum Sharpe ratios relative to the case of the conditionally optimally hedged benchmark assets. For example, when all three FX styles are considered, the Sharpe ratio is increased to 1.62, an increase of 30% relative to the conditionally hedged benchmark. The diversification benefits that derive from the FX style portfolios are economically large and highly significant. Thus, the benefits from the FX styles due to the speculation component are also significant when carefully accounting for the hedging component.

Panel B of Table 18 presents results for when the FX investments are restricted to the G10 currencies. For the carry trade and the FX value strategy, the results are qualitatively similar to the "all currencies" case. Increases in Sharpe ratios are economically large, although they are somewhat lower compared to Panel A. The most notable difference is for FX momentum. Adding the FX momentum strategy which only invests in G10 currencies does not provide significant diversification benefits. This result reflects the facts, documented by Menkhoff, Sarno, Schmeling, and Schrimpf (2012b), that momentum strategies do not perform well for developed market currencies, and that transaction costs due to portfolio re-balancing are large compared

[52]Note that single currency positions are only included for hedging, but not for speculation. If we also allow for speculation, the maximum Sharpe ratio is identical for all hedging schemes by construction and only the changes of the Sharpe ratios differ. Appendix C.4 provides results for including speculation benefits from single currencies in the benchmark. The conclusions on the diversification gains from the FX style investments are unaltered.

to the carry trade.[53] Finally, when we jointly test all three FX styles restricted to the G10 currencies, we can reject equivalence of the Sharpe ratios at the 1% level for the regression-based test and at the 5% level for the SDF-based test. This finding of an improvement in the portfolio allocation holds for all three hedging schemes considered.

Further descriptive statistics of the resulting portfolio characteristics are presented in Table 19 (Panel A). The table also provides the positions (Panel B) of the benchmark allocations and for the case in which the three FX styles are added to the investment universe. Besides the notable effects on the Sharpe ratio, it is notable that adding all three FX styles results in rather weakly skewed return distributions, ranging from -0.07 (all currencies) to -0.08 (G10 currencies).

The benchmark portfolio positions reported in Panel B indicate that global bonds receive a dominant share in all allocations.[54] When the three FX styles are added to the conditional optimal hedged benchmark assets, the optimal positions for the test assets vary from 7% (FX momentum, based on G10 currencies) to 20% (carry trade, based on all currencies). By construction, these weights reflect the pure speculative positions for the three FX styles. How do these positions compare to the FX style hedging positions for our optimized portfolio? We show in Appendix C.1 that the hedging positions in the three FX styles can be calculated analytically as $\widetilde{\boldsymbol{\Psi}}_{RZ;RM} \times \mathbf{w}_R$, where $\widetilde{\boldsymbol{\Psi}}_{RZ;RM}$ is the matrix of risk management demands for the three FX styles, and $\mathbf{w}_R$ is the vector of optimal portfolio weights for global bonds and global stocks. These hedging positions are -4% for the carry trade, -2% for FX momentum, and +2% for FX value (FX styles based on all currencies). These risk-minimizing hedging positions are an order of magnitude lower than the speculative positions.

[53]In Appendix C.4, we repeat the tests before applying our transaction costs adjustment. These results represent an upper bound of diversification benefits. Here, we find significant benefits also for FX momentum at the 5% level for all three hedging schemes. Thus, the momentum signal in currency markets is not too weak, rather, transaction costs eat up a large proportion of the profits.

[54]To conserve space, the table reports only the sum of the positions of the global bond and equity markets.

Table 19: Portfolio Characteristics and Positions

Panel A displays the mean (in percentage points), standard deviation (Std), skewness (Skew), and the Sharpe ratio of the in-sample optimal portfolios in Table 18. The mean is annualized by multiplying by 12, the standard deviation and the Sharpe ratio (SR) are annualized by multiplying by $\sqrt{12}$. Below the column headers "full hedge", "optimal hedge", and "conditional optimal hedge" we report the properties of the benchmark portfolios of global bonds and global stocks using different hedging schemes. Below the column header "conditional optimal + FXS" we present the characteristics of the conditional optimally hedged portfolio when FX styles ("FXS") are added to the benchmark allocation of conditionally hedged assets. Panel B provides the positions of the benchmark assets and FX styles. The sample period is 02/1981 - 12/2011.

Panel A: Characteristics of optimal portfolios						
	FX styles based on all currencies				G10 currencies	
	full hedge	optimal hedge	cond. opt.	cond. + FXS	cond. opt.	cond. + FXS
Mean	5.42	5.51	6.06	5.37	5.69	5.24
Std	4.65	4.59	4.86	3.32	4.68	3.58
SR	1.17	1.20	1.25	1.62	1.21	1.46
Skew	-0.12	-0.09	-0.08	-0.07	-0.06	-0.08
Panel B: Optimal portfolio positions (max. Sharpe ratio)						
Global bonds (Σ)	0.98	0.99	1.04	0.55	1.00	0.64
Global stocks (Σ)	0.02	0.01	-0.04	-0.02	0.00	0.00
Carry trade	0.00	0.00	0.00	0.20	0.00	0.15
FX momentum	0.00	0.00	0.00	0.12	0.00	0.07
FX value	0.00	0.00	0.00	0.15	0.00	0.14

3.5.3 Out-of-Sample Results

Methods. All of the previous mean-variance efficiency tests are based on in-sample estimation results, that is, the optimal positions in each asset are only determined ex post. A far more realistic assessment is to mimick investment decisions in real-time. In this section, we therefore determine asset allocations ex ante and re-examine our previous results. Similar to DeMiguel, Garlappi, and Uppal (2009), we use optimized as well as naive portfolio formation rules. The naive portfolio formation rule allocates an equal fraction to the three "asset classes", global bonds, global stocks, and FX investment styles. The portfolio of global bonds is a weighted average (GDP at PPP) of the bond returns of the individual markets (as in Table 28). The same applies to the global equity portfolio. The set of test assets is either a specific FX style portfolio or an equally weighted portfolio of all three FX styles. The optimized portfolio formation rule is a mean-variance tangency portfolio where we impose short-sales constraints on all assets. In a rolling sample approach, we take the first 120 observations of our sample to calculate portfolio positions and compute the corresponding portfolio return for the following period. Next, we move the rolling window one period forward and repeat the previous steps. This results in a time series of 251 out-of-sample portfolio returns. We first follow this procedure for the benchmark assets and then apply it to the augmented set of assets including the FX styles.[55] We use the same rolling window approach to apply the optimal hedge and the conditional optimal hedge in our out-of-sample setting.

We notice that we do not intend to compare different portfolio formation rules with each other or even to recommend one of them, as each has its specific drawbacks and difficulties.[56] Rather, for a given portfolio formation rule we focus on the comparison (i.e. the change in the Sharpe ratio) between the portfolio containing the (hedged) benchmark assets and the portfolio containing the augmented asset menu. To conduct statistical inference, we apply the delta method to calculate HAC-robust t-values for the change in the Sharpe ratio, as proposed by Ledoit and Wolf (2008).

Results for the Hedging Strategies. First, we present results for the out-of-sample economic significance of the benefits from the various currency hedging strategies. These results are interesting since Campbell, Sunderam, and Viceira (2010) only provide results for the in-sample performance of the different hedging schemes. Table 20 shows that the full hedge leads to significant improvements for all bond markets and most of the stock markets. We also find that, for global bonds, the optimal hedge is actually counterproductive. Our out-of-sample tests indicate

[55] Appendix C.4 provides results for further portfolio rules (see, e.g. DeMiguel, Garlappi, and Uppal, 2009).

[56] DeMiguel, Garlappi, and Uppal (2009) evaluate the out-of-sample performance of the mean-variance model and the naive 1/N (equal weights) rule across several data sets. Overall, they do not find consistently better results from optimal portfolio formation rules compared to a simple 1/N rule.

a deterioration of the risk characteristics compared to the full hedge. For global stocks, the optimal hedge tends to achieve better results than the full hedge in most cases. However, the gains are only statistically significant (5% level) for four out of the eleven equity markets. Finally, the conditional optimal hedge generally does not perform very well out of sample compared to the other hedging strategies. Only for three out of the 22 assets does the conditional optimal hedge provide some benefits against the optimal hedge in the out-of-sample analysis, and these cases are not statistically significant. In sum, the full hedge emerges as a sound alternative to more sophisticated hedging schemes in our out-of-sample setting, coming close to optimal even for global equities.

Table 20: Out-of-Sample Economic Significance of Risk Management Demands

The table shows the standard deviation of global bonds and equities (the benchmark assets) for different hedging schemes in an out-of-sample setting. The Campbell, Sunderam, and Viceira (2010) risk management demands for the optimal hedge and the conditional optimal hedge are calculated using a 120 months rolling window approach. We use the first 120 observations of our sample to calculate the hedging weights for the hedged benchmark asset return in period 121. Next, we move the rolling window one period forward and repeat the previous step, which results in a time-series of out-of-sample hedged benchmark asset returns. The sample period is 02/1981 - 12/2011. All performance results, also for the no hedge and the full hedge schemes, are based on the sample period 02/1991 - 12/2011.

	Standard deviation % $\times\sqrt{12}$				Tests of significance					
	no hedge	full hedge	opt. hedge	cond. opt.	FH vs. NO		OH vs. FH		OH vs. CO	
	(NO)	(FH)	(OH)	(CO)	F-stat	p-val.	F-stat	p-val.	F-stat	p-val.
GDP PPP weighted portfolios										
Global bonds	6.74	4.82	4.87	4.91	1.96	0.00	0.98	0.57	0.98	0.56
Global stocks	15.13	14.23	12.42	12.28	1.13	0.17	1.31	0.02	1.02	0.43
Individual global bond markets										
U.S.		6.45	6.57	6.58			0.96	0.61	1.00	0.51
AU	13.07	6.73	6.85	7.05	3.77	0.00	0.97	0.60	0.94	0.68
CA	9.69	5.85	6.01	6.19	2.74	0.00	0.95	0.66	0.94	0.68
JP	11.65	4.78	4.83	4.88	5.95	0.00	0.98	0.57	0.98	0.56
GE	12.74	5.62	6.08	6.28	5.14	0.00	0.85	0.89	0.94	0.70
NZ	13.45	6.88	7.23	7.48	3.82	0.00	0.90	0.79	0.93	0.71
NO	12.55	7.72	8.04	8.21	2.64	0.00	0.92	0.74	0.96	0.64
SE	13.00	6.51	6.65	6.89	3.99	0.00	0.96	0.63	0.93	0.71
SW	12.80	5.11	5.49	5.62	6.27	0.00	0.87	0.87	0.96	0.64
U.K.	10.72	6.03	6.31	6.45	3.16	0.00	0.91	0.76	0.96	0.63
Individual global stock markets										
U.S.		15.14	14.05	13.70			1.16	0.12	1.05	0.34
AU	20.97	13.50	12.92	12.90	2.42	0.00	1.09	0.24	1.00	0.49
CA	20.28	15.44	13.16	13.11	1.73	0.00	1.38	0.01	1.01	0.48
JP	23.09	21.44	19.58	19.94	1.16	0.12	1.20	0.08	0.96	0.61
GE	20.35	18.93	18.08	18.47	1.16	0.13	1.10	0.23	0.96	0.63
NZ	22.40	17.01	17.17	17.46	1.74	0.00	0.98	0.56	0.97	0.60
NO	27.19	23.45	20.85	21.28	1.34	0.01	1.27	0.03	0.96	0.63
SE	26.65	24.07	22.62	23.19	1.23	0.05	1.13	0.16	0.95	0.65
SW	17.11	16.01	14.80	15.14	1.14	0.15	1.17	0.11	0.96	0.64
U.K.	16.49	14.50	12.96	13.21	1.29	0.02	1.25	0.04	0.96	0.62

Results for the FX Styles. We now proceed to the quantitative out-of-sample (OOS) analysis of the FX styles. Do the diversification benefits of the FX strategies, documented via in-sample tests (in Table 2), also show up in a more realistic out-of-sample setting? In Panel A of Table 21 we compare the OOS Sharpe ratios of naive asset allocations, without and with FX investment styles based on all currencies. When we fully hedge the benchmark assets, the annualized Sharpe ratio of the benchmark is 0.63. The Sharpe ratio increases further to 0.88, 0.83, and 0.84, if the asset menu is augmented by the carry trade, FX momentum, or FX value strategies. Once again, we find additional benefits if we add all three FX styles jointly to the investment universe. The Sharpe ratio is as large as 0.91 in this specification, which represents an improvement of about 30% relative to the fully hedged benchmark. These improvements are statistically significant as can be inferred from the t-values. Results based on other hedging strategies for the benchmark (optimal and conditional optimal) speak the same language. Since the optimal and the conditional optimal hedge tend to perform somewhat worse than the full hedge out-of-sample, the relative improvements from adding the FX styles to the allocation are even somewhat larger (compared to the case of the fully hedged benchmark).

Panel B of Table 21 provides results for the mean-variance optimized portfolio formation rule. Overall, the Sharpe ratios of the portfolios formed by the optimized rule exceed those of the naive portfolios. The diversification benefits from including the FX investment styles, measured by the changes of the Sharpe ratios, are large and highly significant in almost all cases. The main exception is momentum, where the results are only borderline significant.

Table 21: Out-of-Sample Sharpe Ratios: All Currencies

The table reports Sharpe ratios for allocations with and without FX investment styles in an out-of-sample setting. The benchmark allocations ("Bench.") are based on global bonds and global stocks (covering Australia, Canada, Germany, Japan, New Zealand, Norway, Sweden, Switzerland, U.K., U.S.). The test assets are FX investment styles adjusted for transaction costs and are based on a broad set of up to 30 currencies. The test portfolio allocations are based on the benchmark assets augmented with one specific FX investment style or augmented with all three FX investment styles ("+FXS"). For the naive portfolio formation rule in Panel A, we allocate 1/2 to global stocks (GDP at PPP weighted), and 1/2 to global bonds (GDP at PPP weighted). For the augmented portfolios, we weight global bonds, global stocks, and FX investment styles 1/3. The optimization portfolio formation rule in Panel B is a mean-variance tangency portfolio with short-selling constraints. We use the first 120 observations of our sample to calculate the portfolio weights for the return in period 121. Next, we move the rolling window one period forward and repeat the previous step, which results in a time-series of out-of-sample returns of one benchmark and one test portfolio. The 120 months rolling window scheme is also used to find out-of-sample risk management demands (optimal hedge / conditional optimal hedge) for the naive as well as for the optimized portfolio formation rule. We report HAC-robust t-statistics for the difference in Sharpe ratios ($\triangle SR$) between the benchmark allocations and the test portfolio allocations in brackets (Newey and West, 1987 kernel with four lags). The sample period is 02/1981 - 12/2011, all portfolio performance results are based on the sample period 02/1991 - 12/2011.

Benchmark: Global bonds and global stocks							
	full hedge $\widetilde{\mathbf{\Psi}}_{RM} = -\mathbf{1_j}$		optimal hedge $\widetilde{\mathbf{\Psi}}_{RM} = -\mathbf{B_{RX}}$		conditional optimal $\widetilde{\mathbf{\Psi}}_{RM} = [-\mathbf{B_{RX}}, -\mathbf{B_{RZ}}]'$		
	Sharpe ratio	$t_{\triangle SR}$	Sharpe ratio	$t_{\triangle SR}$	Sharpe ratio		$t_{\triangle SR}$
Panel A: Naive portfolio formation (1/3 stocks, 1/3 bonds, 1/3 FX styles)							
Bench.	0.63		0.50		Bench.	+FXS	
Carry trade	0.88	[2.17]	0.84	[2.53]	0.44	0.83	[2.71]
FX momentum	0.83	[1.94]	0.71	[1.97]	0.48	0.71	[2.04]
FX value	0.84	[2.20]	0.72	[2.27]	0.46	0.69	[2.31]
All	0.91	[4.52]	0.82	[4.41]	0.45	0.79	[4.51]
Panel B: Mean-variance optimized portfolio formation (without short-sales)							
Bench.	0.76		0.60		Bench.	+FXS	
Carry trade	1.09	[2.44]	0.89	[1.89]	0.67	0.90	[1.44]
FX momentum	0.87	[1.34]	0.74	[1.81]	0.62	0.75	[1.72]
FX value	0.96	[2.01]	0.81	[2.14]	0.59	0.81	[2.36]
All	1.26	[3.48]	1.05	[3.06]	0.64	1.09	[2.71]

Table 22 reports results for FX styles restricted to the G10 currencies. The changes in Sharpe ratios tend to be smaller in magnitude in all specifications. There are also several cases where adding a single FX investment style portfolio does not result in a significant improvement in the portfolio allocation. When we test for the benefits from adding all three FX styles jointly, however, we can reject equivalence of the Sharpe ratios at the 5% level. This holds for all three hedging schemes and the naive as well as the optimized portfolio formation rule.

Table 22: Out-of-Sample Sharpe Ratios: G10 Currencies

The table reports Sharpe ratios for allocations with and without FX investment styles in an out-of-sample setting. The benchmark allocations ("Bench") are based on global bonds and stocks (covering Australia, Canada, Germany, Japan, New Zealand, Norway, Sweden, Switzerland, U.K., U.S.). The test assets are FX investment styles adjusted for transaction costs and are based on G10 currencies. The test portfolio allocations are based on the benchmark assets augmented with one specific FX investment style or augmented with all three FX investment styles ("+FXS"). For the naive portfolio formation rule in Panel A, we allocate 1/2 to global stocks (GDP at PPP weighted), and 1/2 to global bonds (GDP at PPP weighted). For the augmented portfolios, we weight global bonds, global stocks, and FX investment styles 1/3. The optimization portfolio formation rule in Panel B is a mean-variance tangency portfolio with short-selling constraints. We use the first 120 observations of our sample to calculate the portfolio weights for the return in period 121. Next, we move the rolling window one period forward and repeat the previous step, which results in a time-series of out-of-sample returns of one benchmark and one test portfolio. The 120 months rolling window scheme is also used to find out-of-sample risk management demands (optimal hedge / conditional optimal hedge) for the naive as well as for the optimized portfolio formation rule. We report HAC-robust t-statistics for the difference in Sharpe ratios ($\triangle SR$) between the benchmark allocations and the test portfolio allocations in brackets (Newey and West, 1987 kernel with four lags). The sample period is 02/1981 - 12/2011, all portfolio performance results are based on the sample period 02/1991 - 12/2011.

Benchmark: Global bonds and global stocks							
	full hedge $\widetilde{\mathbf{\Psi}}_{RM} = -\mathbf{1_j}$		optimal hedge $\widetilde{\mathbf{\Psi}}_{RM} = -\mathbf{B_{RX}}$		conditional optimal $\widetilde{\mathbf{\Psi}}_{RM} = [-\mathbf{B_{RX}}, -\mathbf{B_{RZ}}]'$		
	Sharpe ratio	$t_{\triangle SR}$	Sharpe ratio	$t_{\triangle SR}$	Sharpe ratio		$t_{\triangle SR}$
Panel A: Naive portfolio formation (1/3 stocks, 1/3 bonds, 1/3 FX styles)							
Bench.	0.63		0.50		Bench.	+FXS	
Carry trade	0.75	[1.11]	0.73	[1.67]	0.46	0.71	[1.77]
FX momentum	0.67	[0.37]	0.54	[0.36]	0.50	0.54	[0.42]
FX value	0.83	[1.46]	0.69	[1.54]	0.44	0.65	[1.70]
All	0.81	[2.37]	0.72	[2.64]	0.43	0.67	[2.78]
Panel B: Mean-variance optimized portfolio formation (without short-sales)							
Bench.	0.76		0.60		Bench.	+FXS	
Carry trade	0.91	[1.51]	0.77	[1.63]	0.58	0.74	[1.63]
FX momentum	0.78	[0.34]	0.64	[1.08]	0.62	0.65	[0.75]
FX value	0.89	[1.33]	0.72	[1.30]	0.56	0.73	[1.70]
All	1.01	[2.20]	0.86	[2.21]	0.52	0.80	[2.30]

Finally, Table 23 shows the portfolio characteristics (Panel A) and average positions (Panel B) for the out-of-sample allocations.[57] Similar to the in-sample case, these results suggest that

[57]To save space, we report results only for the full hedge and the conditional optimal hedge. Results for the

the improvements in portfolio characteristics go beyond the first two moments. Adding all three FX styles to a traditional asset portfolio also yields better downside risk characteristics as measured by the skewness, the 5% Value at Risk, and the 5% Expected Shortfall. The average positions for the mean-variance optimized portfolios are similar to those reported earlier.

Table 23: Out-of-Sample Portfolio Characteristics and Positions

The table reports portfolio characteristics and average positions in benchmark assets ("Bench") and the three FX investment styles ("+FXS") of the out-of-sample allocations in Table 21. VaR is the Value at Risk. ES is the Expected Shortfall and is based on the historical return distribution. Returns are annualized by multiplying with 12, the standard deviation is annualized by multiplying with $\sqrt{12}$.

Panel A: Characteristics of out-of-sample portfolios; FX styles are based on all currencies								
	Naive portfolios				Mean-variance without short sales			
	full hedge	full + FXS	cond. opt.	cond. + FXS	full opt.	full + FXS	cond. opt.	cond. + FXS
Mean	4.49	4.72	3.02	3.74	3.50	4.14	3.11	3.69
Std	7.17	5.20	6.71	4.74	4.59	3.29	4.86	3.40
SR	0.63	0.91	0.45	0.79	0.76	1.26	0.64	1.08
Skew	-0.66	-0.51	0.57	0.57	-0.38	-0.05	-0.45	-0.27
VaR 5%	39.11	27.81	33.15	22.85	20.51	16.92	23.43	16.56
ES 5%	54.64	37.07	43.85	28.10	32.93	22.11	34.91	24.62
Panel B: Out-of-sample portfolio positions								
Global bonds (Σ)	0.50	0.33	0.50	0.33	0.86	0.52	0.85	0.49
Global stocks (Σ)	0.50	0.33	0.50	0.33	0.14	0.05	0.15	0.09
Carry trade	0.00	0.11	0.00	0.11	0.00	0.19	0.00	0.18
FX momentum	0.00	0.11	0.00	0.11	0.00	0.10	0.00	0.09
FX value	0.00	0.11	0.00	0.11	0.00	0.15	0.00	0.16

3.5.4 Summary of Main Results

Overall, we find evidence of quantitatively large and significant diversification benefits for almost all of the three FX styles. The results generally hold when the FX styles are based on all currencies or based on G10 currencies, and when the benchmark assets are fully hedged, optimally

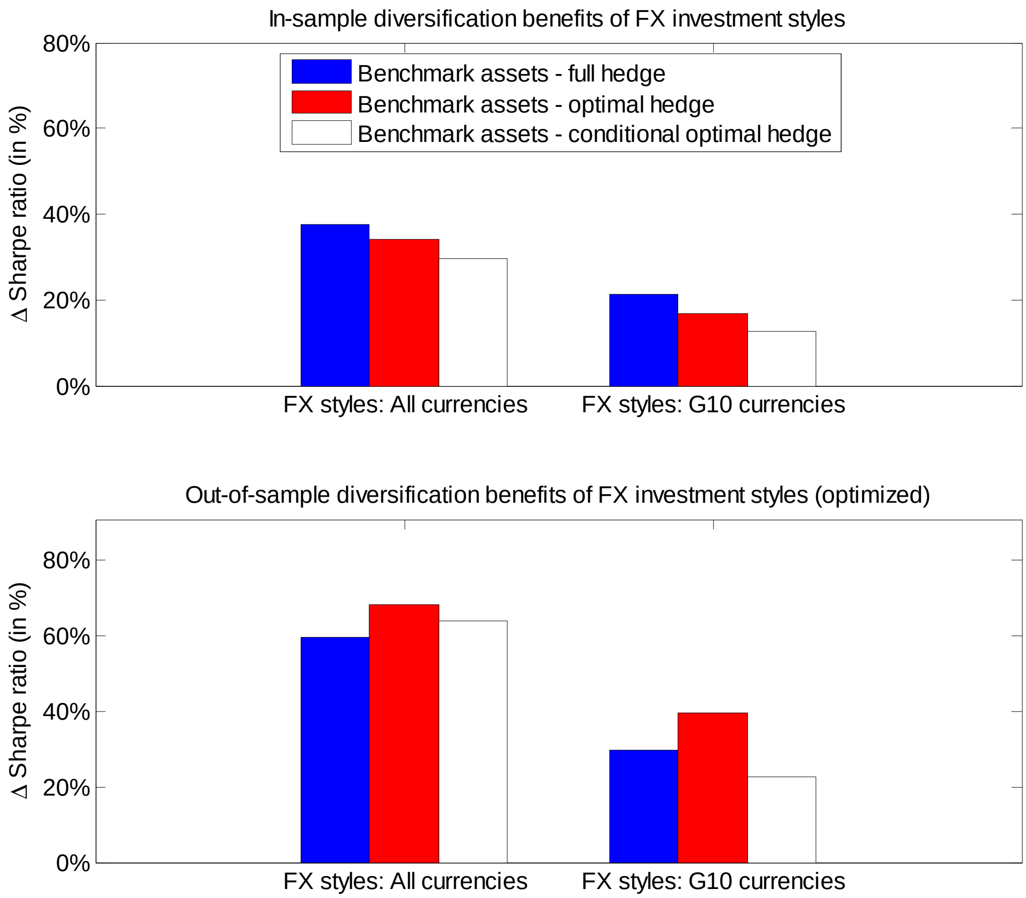
optimal hedge can be provided by the authors upon request.

hedged, or conditionally optimally hedged. The most notable exception is FX momentum, which only provides diversification benefits if investments in emerging market currencies are available to the investor.

The diversification benefits from FX style investing also show up in an out-of-sample evaluation mimicking investor behavior in real-time. Not surprisingly, the levels of Sharpe ratios are lower out-of-sample, but also more realistic. Importantly, the differences in Sharpe ratios between the benchmark allocations and the allocations including FX styles are still economically large and in most specifications statistically significant. Interestingly, the best results are found for a combination of all three FX investment styles. Figure 9 summarizes our baseline results and documents how a portfolio of global bonds and equities can be improved (measured by the increase of the Sharpe ratio in %) by employing all three FX trading strategies. Based on these major findings, we explore our baseline results from several other angles in the following section.

Figure 9: Overview on Baseline Results

The figure summarizes the increases of the Sharpe ratios (in %) when all three FX investment styles are added to a benchmark of global bonds and global stocks, given the different specifications of the Tables 2, 4 and 5. Out-of-sample results are based on the mean-variance portfolio formation rule with imposing short-sales constraints.



3.6 Further Results

3.6.1 Value and Momentum Everywhere

To further assess the diversification potential of FX investment strategies, we thoroughly tested for the robustness of the results with respect to changes in the set of benchmark assets. Our baseline results (reported in Tables 18) were based on buy and hold positions in the benchmark assets. Now, we consider a set of popular managed portfolios, namely value and momentum portfolios for other asset classes.[58]

We collect the value and momentum "everywhere" portfolios of Asness, Moskowitz, and Pedersen (2012) to check whether FX investment styles provide benefits that go beyond those of value and momentum strategies in other asset classes. We choose their rank-based portfolios which are constructed in a similar way as our FX styles. The benchmark consists of portfolios of global stocks and global bonds, and value and momentum portfolios for U.S., U.K., European, and Japanese equities, country stock market indices, and commodities ($K=2+2\times6$).[59] The sample for all assets covers the period from 07/1981 to 06/2010.

Table 24 provides the results for the mean-variance efficiency tests. We find that the FX investment strategies also yield diversification benefits in this setting where the benchmark assets are managed portfolios. For example, if the benchmark assets are optimally hedged, the Sharpe ratio increases from 2.01 to 2.28 (2.12, for G10 currencies) if we add all three FX style portfolios. This change in return per unit of risk is significant at the 1% level for the FX styles based on all currencies. For the FX styles restricted to G10 currencies, however, only the regression-based test indicates significance at the 10% level. Note, however, that unlike our FX style portfolios, Asness, Moskowitz, and Pedersen (2012) do not account for transaction costs due to portfolio rebalancing.[60] Appendix C.4 provides additional results for FX investment styles without taking transaction costs into account, since these test assets are more in line with the benchmark assets. We find larger increases of Sharpe ratios and considerably lower p-values for the tests of mean-variance efficiency. For example, all three FX styles based on G10 currencies increase the Sharpe ratio to 2.14 which is significant at the 5% level according to the regression-based test. Thus,

[58] Fama and French (1998) and Rouwenhorst (1998) document value and momentum premia in international stock markets, and Eun and Lee (2010a) find large diversification benefits for international style-based equity investing. Value or momentum strategies are also found profitable for country stock market indices, Asness, Liew, and Stevens (2000) and Bhojraj and Swaminathan (2006), and for commodities, Gorton, Hayashi, and Rouwenhorst (2012).

[59] Descriptive statistics for these alternative benchmark assets are provided in Appendix C.4. The European value and momentum portfolios and the country indices portfolios are exposed to unknown long-short positions in multiple currencies. Thus, it is not feasible to apply the full hedge to all benchmark assets.

[60] Sadka and Korajczyk (2004) show that equity momentum returns are substantially lower when realistic transaction cost adjustments are considered.

these results on the diversification benefits of FX investment strategies can be considered to be conservative.[61]

Table 24: FX Styles and Value and Momentum "Everywhere"

The table displays mean-variance efficiency tests as described in Table 18, except that the benchmark includes one global bonds portfolio (GDP at PPP weighted), one global stocks portfolio (GDP at PPP weighted), as well as twelve (rank-based) value and momentum portfolios across countries and asset classes provided by Asness, Moskowitz, and Pedersen (2012). These portfolios cover one value and one momentum strategy for the stock markets of the U.S., the U.K., Europe, Japan, country indices, and commodities (K=2+12). The FX styles are based on all currencies in Panel A and they are based on the G10 currencies in Panel B. The FX styles account for transaction costs due to portfolio re-balancing. The benchmark assets by Asness, Moskowitz, and Pedersen (2012) do not account for any transaction costs. The sample period is 07/1981 - 06/2010.

Benchmark: Value and momentum "everywhere" (not adjusted for transaction costs)							
	optimal hedge $\widetilde{\mathbf{\Psi}}_{RM} = -\mathbf{B_{RX}}$			conditional optimal $\widetilde{\mathbf{\Psi}}_{RM} = [-\mathbf{B_{RX}}, -\mathbf{B_{RZ}}]'$			
	W	SDF	Sharpe	W	SDF	Sharpe	
	p-value		ratio	p-value		ratio	
Panel A: FX styles based on all currencies (adjusted for transaction costs)							
		Bench.:	2.01			Bench.	+FXS
Carry trade	0.000	0.002	2.20	0.001	0.007	2.07	2.21
FX momentum	0.011	0.019	2.07	0.012	0.020	2.01	2.07
FX value	0.051	0.081	2.07	0.020	0.040	1.99	2.06
All	0.000	0.002	2.28	0.000	0.005	2.04	2.29
Panel B: FX styles based on G10 currencies (adjusted for transaction costs)							
		Bench.:	2.01			Bench.	+FXS
Carry trade	0.060	0.099	2.06	0.026	0.057	1.98	2.04
FX momentum	0.405	0.365	2.02	0.267	0.233	2.01	2.02
FX value	0.037	0.059	2.08	0.037	0.047	2.01	2.08
All	0.082	0.258	2.12	0.049	0.164	1.97	2.10

[61] Besides transaction costs, there is another reason why our results for the benefits of FX styles are conservative. Long-short portfolios are easy to build in FX forward markets but more difficult to implement in equity markets. However, Eun and Lee (2010a) report that the achievable Sharpe ratio of the long only leg of global equity styles is considerably smaller. They find that after including short-sales constraints in global equity markets, the maximum Sharpe ratio of global equity styles (i.e. our benchmark here) is reduced by approximately two thirds.

3.6.2 Accounting for Skewed Returns

It is well known that some of the FX investment strategies, especially the carry trade, are prone to occasional large losses, that is, the return distribution is negatively skewed (see e.g. Gyntelberg and Remolona, 2007; Brunnermeier, Nagel, and Pedersen, 2009). The standard mean-variance framework ignores this characteristic. In the following, we consider three approaches to account for negative skewness. The first two approaches change the construction of the FX investment portfolios to reduce the negative skewness. The third approach considers an alternative test methodology which explicitly accounts for skewed return distributions.

Options. We first study a modified carry trade strategy where the downside risk is insured using currency options as in Jurek (2009) and Burnside (2011). For every long position, an options-based hedging portfolio buys an equal amount of put options such that the investor is compensated for large losses. Similarly, for every short position, an options-based hedging portfolio buys an equal amount of call options. This hedge for downside risk is rather extreme. By construction, given a specific threshold, the investor insures all of the downside risk of the carry trade. We obtain options data from J.P. Morgan for our set of G10 currencies for a limited sample period from 07/1997 to 12/2011. Further details on these data and the construction of the options-based hedge are provided in the Appendix C.1 to this paper.

Panel A of Table 25 reports the characteristics of the carry trade (based on G10 currencies) with and without options-based downside risk insurance. By construction, the skewness of carry returns is much reduced (in absolute terms) to -0.11 (+0.07) in case of the 10 delta (25 delta) hedge. Burnside (2011) report that the costs of these hedges are low in terms of Sharpe ratios, in particular compared to a similar strategy in stock markets. In line with their results, we find that the Sharpe ratio of the carry trade 10 delta (25 delta) hedge is still as large as 0.47 (0.40) compared to 0.55 for the case without downside risk insurance.

Panel B of Table 25 provides tests for mean-variance efficiency for the options-hedged carry strategies. The benchmark assets are global bonds and equities as in our baseline results. Adding the baseline carry trade increases the Sharpe ratio for the fully hedged benchmark assets from 1.55 to 1.66 for this shorter subsample. This increase is significant at the 5% level for both tests. The increase in the Sharpe ratio is 1.62 if the carry trade is hedged using 10 delta options ("far out-of-the-money"), a significant increase at the 10% level for the regression-based test but not for the SDF-based test. Finally, the increase in the Sharpe ratio is 1.60 if the carry trade is hedged using 25 delta options ("out-of-the-money") and (marginally) no longer significant. The results for the optimally hedged and conditional optimally hedged benchmark assets are similar. Thus, insuring all downside risk inherent in the carry trade with options seems to be costly in that it tends to reduce the benefits of diversification.[62]

[62]Note, however, that the return distributions of the benchmark assets also exhibit a fair amount of negative skewness. Hence, to level the playing field these assets ought (in principle) be hedged for downside risk as well. Burnside (2011) show that a similar downside hedging strategy for U.S. stocks reduces returns from 6.87% to -4.79% (p.a.). Downside risk protection for U.S. equities is therefore far more costly than for the carry trade.

Table 25: Carry Trade Hedged for Downside Risk

Panel A presents the characteristics of the carry trade (based on G10 currencies) hedged with put and call options. The hedging portfolios are constructed using put and call options such that negative skewness cancels out as described by Burnside (2011). Option data are based on implied volatility quotes and are from J.P. Morgan. Quotes are available for 25 delta ("out-of-the-money") and 10 delta ("far out-of-the-money") call and put options. The mean is annualized by multiplying by 12, the standard deviation and the Sharpe ratio (SR) are annualized by multiplying by $\sqrt{12}$. Panel B displays mean-variance efficiency tests for the hedged carry trade strategies as described in Table 18. The benchmark assets are global bonds and global stocks (K=20). The sample period is 07/1997 - 12/2011.

Panel A: Risk and return characteristics
Carry trade hedged with put and call options (G10 currencies, 07/1997 - 12/2011)

	Mean	Std	Skew	SR
Carry trade	4.48	8.11	-0.43	0.55
Carry + 10 delta	3.45	7.27	-0.11	0.47
Carry + 25 delta	2.68	6.69	0.07	0.40

Panel B: Mean-variance efficiency tests
Benchmark: Global bonds and global stocks (07/1997 - 12/2011)

	full hedge $\widetilde{\Psi}_{RM} = -\mathbf{1_j}$			optimal hedge $\widetilde{\Psi}_{RM} = -\mathbf{B_{RX}}$			conditional optimal $\widetilde{\Psi}_{RM} = [-\mathbf{B_{RX}}, -\mathbf{B_{RZ}}]'$			
	W	SDF	Sharpe	W	SDF	Sharpe	W	SDF	Sharpe	
	p-value		ratio	p-value		ratio	p-value		ratio	
		Bench.:	1.55		Bench.:	1.59			Bench.	+FXS
Carry trade	0.033	0.047	1.66	0.047	0.080	1.69	0.050	0.081	1.66	1.75
Carry + 10 delta	0.094	0.110	1.62	0.098	0.099	1.67	0.103	0.098	1.61	1.67
Carry + 25 delta	0.113	0.119	1.60	0.118	0.107	1.65	0.151	0.135	1.61	1.66

FX strategy diversification. As another possibility we consider whether exposures to higher moments can be diversified to some extent by using combinations of the three FX strategies. Thus, for a moment, we ignore the further diversification potential from global bonds or global stocks. As a first step, we calculate the model-free moments implied by our options data for each of the G10 currencies, as proposed by Bakshi, Kapadia, and Madan (2003). For each FX investment style, we then compute a weighted average of the model-free implied moments of the different currencies. The individual weights are determined according to our FX style portfolio formation approach. In this way we obtain option-based "ex ante" gauges of FX style risk with respect to volatility, skewness, and kurtosis.

The interpretation of these measures is simple. Our FX style portfolios are one USD long in a basket of currencies, and one USD short in another basket of currencies. Thus, if the option-based measure for a particular moment is positive, the particular FX style has the tendency to be long in currencies which are more exposed to that specific moment, and has the tendency to be short in currencies which are less exposed to the specific moment.

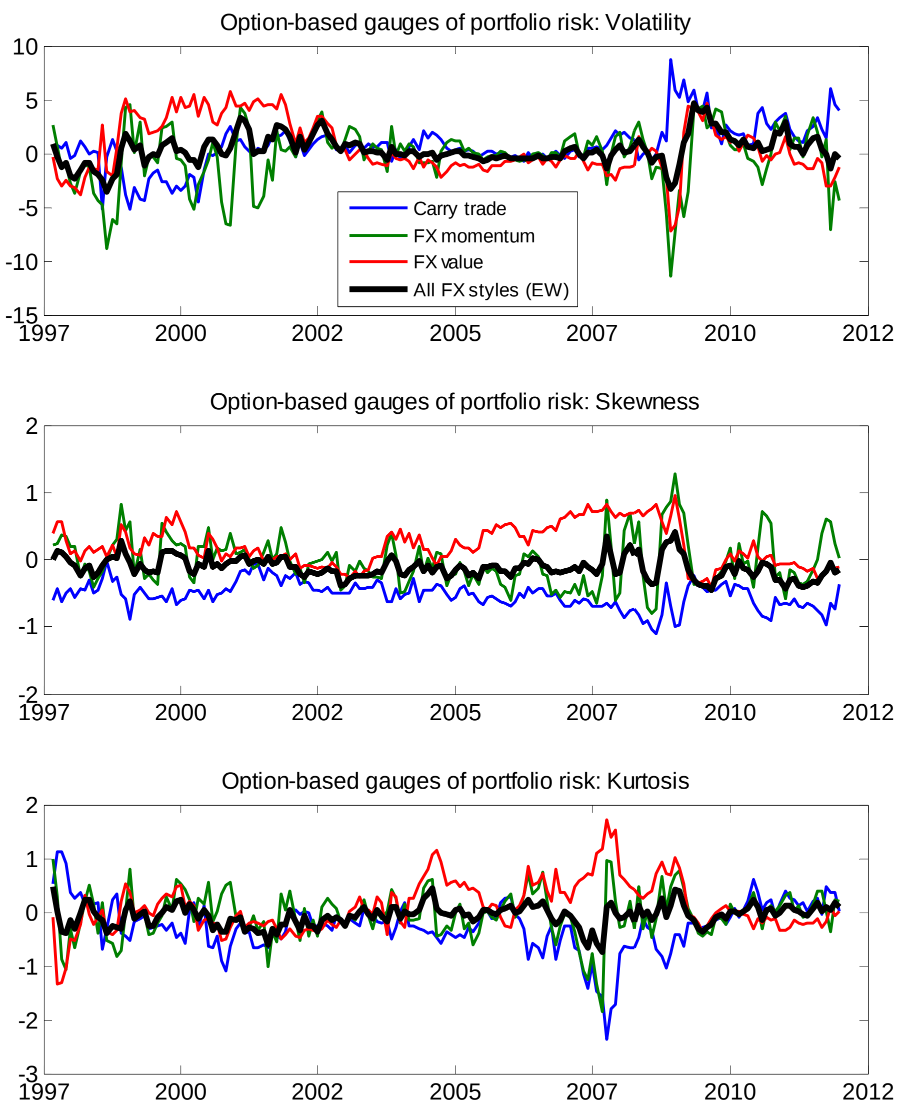
The time-series of the option-based gauges of risk are depicted in Figure 10, yielding some interesting additional insights. In line with the earlier literature, we find that the carry trade investment currencies are negatively exposed to skewness. In contrast, FX value long currencies tend to be positively exposed to skewness. There is no clear pattern for the FX momentum currencies. Importantly, the figure also reports the exposures to higher moments for an equally weighted strategy of all three FX styles. The average currency of a strategy that invests equal amounts in all three FX styles is never highly exposed in one or the other direction. Thus, a strategy diversified across all three FX strategies appears to be fairly neutral with respect to skewness. We find a similar diversification pattern for the option-based measure for volatility and kurtosis.

Panel A of Table 26 reports the return and risk characteristics of a simple equally weighted portfolio of the FX strategies. The FX styles are based on all currencies or the G10 currencies and we can make use of the full sample period from 02/1981 to 12/2011. The first three strategies are equally weighted across two FX styles (Carry trade and FX momentum, Carry trade and FX value, FX momentum and FX value). The fourth strategy is equally weighted across all three FX styles. A strategy which combines all three signals is also studied by Jorda and Taylor (2012).[63] We find that the strategy combining FX momentum and FX value is positively skewed and provides a Sharpe ratio of 0.71 (0.48, when restricting to G10 currencies). The strategy combining all three FX styles shows a small negative skewness of -0.03 (-0.32, G10) and yields

[63]Jorda and Taylor (2012) provide evidence that FX trading strategies not only based on carry, but augmented by past changes of spot rates and real exchange rates, show less skew and large Sharpe ratios.

Figure 10: Option-based Gauges of FX Style Risk

The table shows the (ex ante) model free implied volatility, skewness, and kurtosis of the G10 currencies weighted by the portfolio positions of three FX styles. The model free implied moments for each of the G10 currencies are calculated using option prices as proposed by Bakshi, Kapadia, and Madan (2003). The sample period is from 07/1997 to 12/2011.



a Sharpe ratio of 1.00 (0.67, G10).

Panel B of Table 26 shows mean-variance efficiency tests for the diversified FX strategies. The results are notable, in particular for the strategies with low exposures to negative skewness. For example, combining a carry with an FX momentum strategy results in an increase of the Sharpe ratio from 1.17 to 1.52. Also the combination of all three FX styles provides significant diversification benefits. The Sharpe ratio increases from 1.17 to 1.57, a rise which is highly significant. The results are robust to the different hedging schemes for the benchmark assets, and restrictions of the FX investment universe. Relative to single FX strategies, diversified strategies seem to be especially beneficial for the smaller set of G10 currencies. These findings suggest that strategy diversification potentially offers improvements in the portfolio allocation in the mean-variances sense, while at the same time also helping to avoid the negative skewness associated with (some) single FX strategies.

Table 26: Diversified FX Investment Styles

Panel A shows the characteristics of equally weighted FX style portfolios. The mean is annualized by multiplying by 12, the standard deviation and the Sharpe ratio (SR) are annualized by multiplying by $\sqrt{12}$. Panel B displays mean-variance efficiency tests as described in Table 18. The benchmark assets are global bonds and global stocks (K=20). The sample period is 02/1981 - 12/2011.

Panel A: Risk and return characteristics
Equally weighted carry trade (C), FX momentum (M) and FX value (V) strategies

	all currencies				G10 currencies			
	Mean	Std	Skew	SR	Mean	Std	Skew	SR
C+M (EW)	4.73	5.39	-0.01	0.88	3.31	6.46	-0.47	0.51
C+V (EW)	4.74	5.16	-0.88	0.92	4.45	6.17	-0.96	0.72
M+V (EW)	3.81	5.34	0.23	0.71	3.05	6.33	0.53	0.48
C+M+V (EW)	4.43	4.40	-0.03	1.00	3.60	5.39	-0.32	0.67

Panel B: Mean-variance efficiency tests
Benchmark: Global bonds and global stocks

	full hedge $\widetilde{\mathbf{\Psi}}_{RM} = -\mathbf{1_j}$			optimal hedge $\widetilde{\mathbf{\Psi}}_{RM} = -\mathbf{B_{RX}}$			conditional optimal $\widetilde{\mathbf{\Psi}}_{RM} = [-\mathbf{B_{RX}}, -\mathbf{B_{RZ}}]'$			
	W	SDF	Sharpe	W	SDF	Sharpe	W	SDF	Sharpe	
	p-value		ratio	p-value		ratio	p-value		ratio	
FX styles based on all currencies										
		Bench.:	1.17		Bench.:	1.20			Bench.	+FXS
C+M (EW)	0.000	0.000	1.52	0.000	0.000	1.54	0.000	0.000	1.27	1.54
C+V (EW)	0.000	0.000	1.52	0.000	0.001	1.52	0.000	0.001	1.22	1.52
M+V (EW)	0.000	0.000	1.35	0.000	0.000	1.39	0.000	0.000	1.19	1.39
C+M+V (EW)	0.000	0.000	1.57	0.000	0.000	1.59	0.000	0.000	1.23	1.59
FX styles based on G10 currencies										
		Bench.:	1.17		Bench.:	1.20			Bench.	+FXS
C+M (EW)	0.002	0.004	1.31	0.005	0.008	1.32	0.007	0.012	1.21	1.32
C+V (EW)	0.000	0.001	1.41	0.000	0.005	1.39	0.000	0.005	1.19	1.39
M+V (EW)	0.019	0.019	1.25	0.019	0.019	1.28	0.015	0.014	1.19	1.28
C+M+V (EW)	0.001	0.001	1.36	0.001	0.003	1.37	0.001	0.003	1.20	1.37

Stochastic dominance tests. As a third approach to account for skewness, we consider a testing method which explicitly accounts for higher moments when testing for improvements in the asset allocation. This framework, which relies on the stochastic dominance criterion, is based on less restrictive assumptions compared to the traditional mean-variance framework and allows to take skewness-loving preferences of investors into account. In particular, a portfolio is second-order stochastic dominance (SSD) efficient if it is optimal for a non-satiable and risk-averse investor, and it is third-order stochastic dominance (TSD) efficient if it is optimal for a non-satiable, risk-averse and skewness-loving investor (Levy, 2006). Our tests draw on the multivariate SSD and TSD tests as proposed by Post and Versijp (2007). In contrast to the other two approaches discussed before, the multivariate stochastic dominance tests compare the return distributions at the portfolio level, and thus, also account for skewness in the benchmark assets.

To conserve space we report the results of the tests and the methodological details in Appendix C.3 of the paper. Overall, our findings on stochastic dominance tests imply that also investors who dislike negatively skewed return distributions would prefer to augment global bond and global stock portfolios with FX styles. Results in Appendix C.3 show that the TSD pricing errors (i.e., "TSD alphas") are economically large and significant for all three FX styles. The p-value for a test statistic of TSD efficiency of the benchmark assets, which is similar in spirit to the mean-variance efficiency tests, is below 1%. Also tests for the less restrictive SSD efficiency of the benchmark assets indicate significant improvements of FX style investments at the 1% level.

3.6.3 Market States

We now examine some dynamic features of diversification, in particular if the benefits depend on the state of the market. Using two different measures of market states, we conduct *conditional* tests of our *out-of-sample* portfolio returns. The first market state measure draws on the work by Cooper, Gutierrez, and Hameed (2004). They define UP markets as months following high 36-month returns of the stock market, and DOWN markets as months following low 36-month returns of the stock market. We use GDP-weighted global stocks as our "stock market", and divide our time t observations into three equally large UP, FLAT, and DOWN market state categories conditional on the time $t-1$ past 36-month return of the stock market. The second measure of market states relies on a macroeconomic variable. Portfolio returns are sorted conditional on the previous month's global industrial production growth (year-on-year).[64] We find that both measures result in qualitatively similar but not identical classifications of UP-FLAT-DOWN market states. Figure 11 indicates that we have roughly two UP and three DOWN market state periods in our sample from 02/1991 to 12/2011 including the dot-com boom and bust as well as the recent financial crisis. The state variables also correspond reasonably well to U.S. recession periods as defined by NBER.

Table 27 shows out-of-sample Sharpe ratios, as described in Table 21, but conditional on the market state. The benchmark portfolio consists of global bonds and equities (fully hedged), and the test portfolio is the benchmark portfolio augmented with the carry trade, FX momentum, FX value, or all three FX investment styles based on all currencies. The augmented portfolios generally show increased Sharpe ratios during all three market states. However, for FX value, the Sharpe ratio decreases during UP market states (statistically insignificant), when we use global industrial production as the conditioning variable.

[64]We collect industrial production as reported by FRED, St. Louis FED, for the same countries as for the stock market except for Canada, Norway, Switzerland, and U.K., as there are no monthly data available for these countries. The global industrial production measure is simply GDP-weighted for the available countries.

Figure 11: Market States

The figure shows how the sample of out-of-sample returns is split into UP, FLAT, or DOWN market states. We use two variables to define market states, one financial and one macroeconomic, and form equally large subsamples based on the one month lagged realization of the variable. Following Cooper, Gutierrez, and Hameed (2004), the first variable is the 36-month cumulative global stock market return, where UP markets are defined based on the highest 1/3 realizations and DOWN markets based on the lowest 1/3 realizations. The global stock market return is GDP at PPP weighted and fully currency hedged. As macroeconomic variable, we use the year on year change of global industrial production (GDP at PPP weighted). The sample period is 02/1991 - 12/2011.

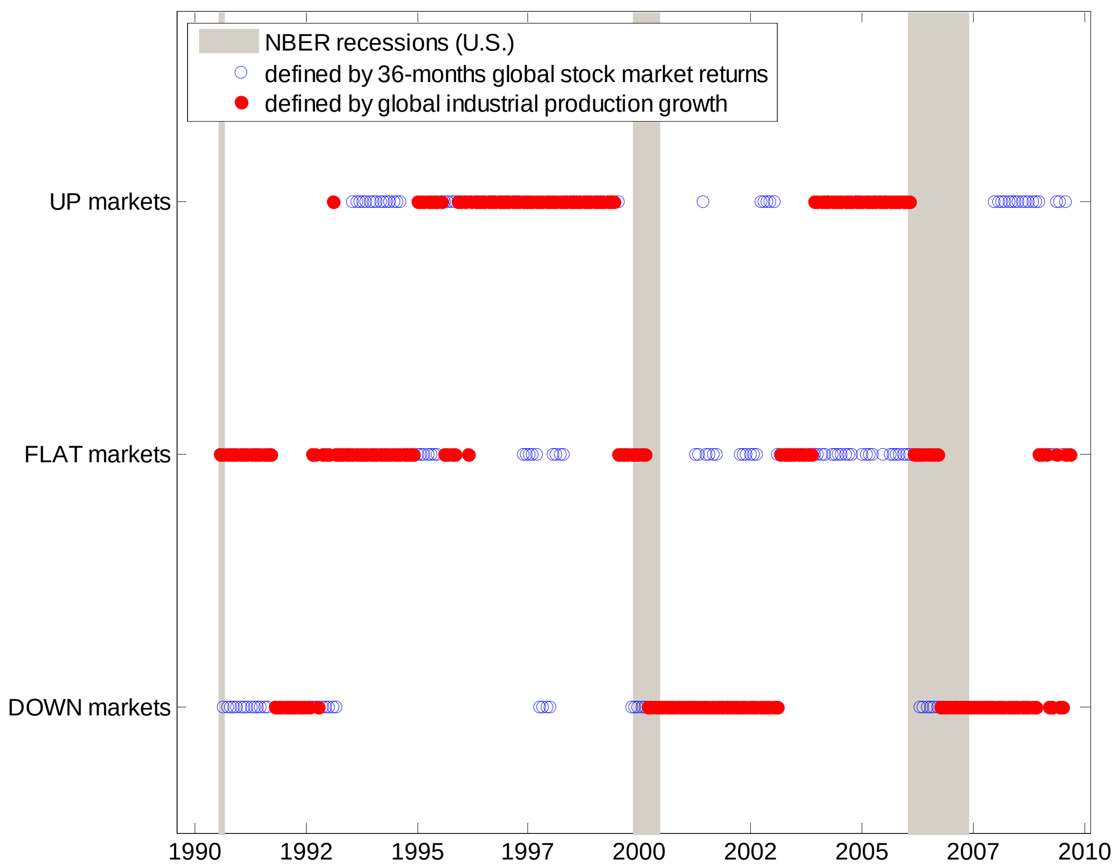


Table 27: Market States and Out-of-Sample Sharpe Ratios

We split our out-of-sample asset allocations covering global bonds and global stocks with and without FX investment styles (as described in Table 21) into three equally large subsamples of UP, FLAT, and DOWN market states. First, following Cooper, Gutierrez, and Hameed (2004), we define UP markets following months with the highest cumulative 36-month global stock market returns. DOWN markets are defined as months following the lowest cumulative 36-month global stock market returns. FLAT markets are defined as all other observations. Second, we use the previous month's Year-on-Year (YoY) change of global industrial production (IP) as a macroeconomic conditioning variable to define market states. The resulting assignment of months to market states for both conditioning variables is shown in Figure 11. We report HAC t-statistics for the difference in Sharpe ratios ($\triangle SR$) between the benchmark allocations and the test portfolio allocations in brackets. The sample period is 02/1981 - 12/2011. We use the first 120 observations to find the initial portfolio weights for the optimized portfolios. All portfolio performance results, also for naive portfolios, are based on the sample period 02/1991 - 12/2011 (251 observations, 84 (83) for each market state).

	Benchmark: Global bonds and global stocks (fully hedged)							
	36-month stock market returns				Global IP growth (YoY)			
	naive		mv-ssc		naive		mv-ssc	
	Sharpe ratio	$t_{\triangle SR}$	Sharpe ratio	$t_{\triangle SR}$	Sharpe ratio	$t_{\triangle SR}$	Sharpe ratio	$t_{\triangle SR}$
Panel A:	UP markets							
Bench. Sharpe ratio:	0.80		0.29		1.06		1.13	
Carry trade	1.04	[1.30]	0.62	[1.70]	1.09	[0.14]	1.35	[1.13]
FX momentum	1.04	[1.32]	0.55	[1.54]	1.22	[0.79]	1.21	[0.51]
FX value	0.86	[0.42]	0.36	[0.38]	0.98	[-0.42]	1.04	[-0.48]
All	1.06	[2.59]	0.82	[2.06]	1.21	[1.51]	1.42	[1.16]
Panel B:	FLAT markets							
Bench. Sharpe ratio:	0.58		0.95		0.85		0.50	
Carry trade	0.93	[1.75]	1.32	[1.70]	1.32	[2.66]	0.87	[2.18]
FX momentum	0.63	[0.27]	0.98	[0.30]	1.03	[1.23]	0.60	[1.11]
FX value	0.75	[0.90]	1.07	[0.85]	1.13	[1.72]	0.80	[1.96]
All	0.84	[2.24]	1.32	[1.63]	1.22	[3.68]	1.04	[2.46]
	DOWN markets							
Bench. Sharpe ratio:	0.53		1.15		0.13		0.67	
Carry trade	0.68	[1.06]	1.44	[1.24]	0.34	[1.28]	1.11	[1.63]
FX momentum	0.78	[1.53]	1.18	[0.13]	0.29	[0.84]	0.82	[0.86]
FX value	0.92	[2.74]	1.50	[1.99]	0.53	[2.77]	1.06	[2.51]
All	0.84	[3.47]	1.76	[2.34]	0.41	[2.83]	1.42	[2.57]

When we focus on the carry trade, there seems to be a hat-shaped pattern in t-values. They are somewhat larger during FLAT market states, and lower during UP and DOWN market states. For FX value, we find the largest t-values during the DOWN market states. There is no obvious pattern for FX momentum. Interestingly, adding all three FX investment styles tends to provide diversification benefits in all three market states. However, we have to note that statistical power should be somewhat smaller in these tests given the small sample size for the market state sub-sample analysis.[65]

3.7 Conclusion

In the past, the existing literature on international diversification often did not assign a high priority to the treatment of the foreign exchange component of international investing. Prior work that does take the FX component seriously mostly focuses on how to eliminate FX risk using various hedging strategies, or how to minimize the risk in global bond or equity portfolios using positions in foreign currencies (see, e.g. Glen and Jorion, 1993; Campbell, Sunderam, and Viceira, 2010).

In this paper, we go beyond the benefits from hedging to shed more light on the speculative component of currency investments. We study the implications of FX investment styles – such as carry trades and widely practiced strategies known as FX momentum, and FX value – for optimal portfolio choice. These strategies are known to be profitable when considered in isolation (see, e.g. Ang and Chen, 2010; Asness, Moskowitz, and Pedersen, 2012; Burnside, Eichenbaum, and Rebelo, 2011; Lustig and Verdelhan, 2007; Menkhoff, Sarno, Schmeling, and Schrimpf, 2012b). But do they provide diversification benefits when the investor already has access to a well-diversified global bond and equity portfolio? To answer this question, we study whether investments in these strategies significantly shift the mean-variance frontier when a benchmark allocation consisting of global bonds and stocks is augmented by FX style investments. Importantly, the benchmark assets are thoroughly hedged in order to dissect the diversification benefits into those deriving from hedging FX risk and those stemming from speculative

[65] We have 251 out-of-sample observations, thus resulting in 84 (83) observations for each market state category.

FX positions.

Our results suggest that style-based FX investments generate significant improvements in the asset allocation. These findings hold after taking into account transaction costs and when controlling for the FX risk inherent in the benchmark assets. Importantly, these results are also confirmed in an extensive out-of-sample experiment with different portfolio formation rules mimicking investor decisions in real-time.

Chapter 4

4 Securitized Real Estate and the Veiling Glare from Currency Risk

This chapter is coauthored by Felix Schindler, and is an earlier Working Paper version of Securitized Real Estate and the Veiling Glare from Currency Risk published in the *Journal of International Money and Finance* 2012, 31, 1851-1866.

4.1 Introduction

Do investments in international securitized real estate markets make a statistically significant contribution to an internationally diversified mixed-asset portfolio, and does currency risk exposure have an impact on these results? These questions have become more and more popular for both private and institutional investors for several reasons. First, (securitized) real estate has been a fast-growing asset class around the world during the last decades. In many countries, REIT legislation has been introduced, improving the institutional framework and legal setting of real estate companies; both the number of listed real estate companies and their market capitalization have increased tremendously, coverage by analysts and investors has augmented, and therefore, securitized real estate now cumulatively offers a suitable opportunity to overcome the drawbacks of investments in direct real estate. Second, many studies show that cross-country diversification benefits for pure stock portfolios have been decreasing over time due to increasing financial integration of global stock markets. Thus, investors are looking for other assets, such as real estate, to provide diversification benefits. Third, recent contributions by Lustig and Verdelhan (2007), Lustig, Roussanov, and Verdelhan (2011), and Verdelhan (2011) provide evidence for systematic risk in exchange rates.[66] If currency risk is not independent and random, buying

[66] Where Lustig and Verdelhan (2007) and Lustig, Roussanov, and Verdelhan (2011) focus on portfolios of exchange rates, Verdelhan (2011) focuses on bilateral exchange rates. Verdelhan finds that "two economically motivated factors" account for 20% to 90% of exchange rate movements.

bonds, stocks, and real estate in different currencies does not lead currency exposure to diversify away in international portfolios (see Verdelhan (2011), p. 6). Consequently, the portfolio diversification benefits from an asset class like international real estate may be masked by common movements in currency risk premia, and thus may lead to an underestimation of the true benefits. Therefore, given the mentioned facts and the increased relevance of real estate investments in the recent past, the central questions raised above present the guideline and motivation for our analysis.

To our knowledge, the paper at hand is the first study applying mean-variance efficiency tests to such a broad range of markets and assets, covering a time period of more than 25 years, while simultaneously considering currency risk, which we regard as an important contribution to the existing literature of international portfolio diversification where currency risk is often neglected even it is a substantial part of risk.

Compared to previous studies applying spanning tests as discussed below, our analysis extends the investment universe from international bond and stock markets to international securitized real estate markets and explicitly considers the impact of currency risk. For the empirical analysis, we use monthly data on bond, stock, and real estate markets from ten countries covering the period from 1984 to 2010. The markets considered, which are located in Asia, Australia, Europe, and North America, cover a large portion of market capitalization in global stock and real estate markets and can therefore be considered representative.

The statistical significance of the diversification benefits is analyzed by regression-based spanning tests for the complete efficient frontier and intersection tests for the global minimum variance portfolio as well as the tangency portfolio as suggested by Huberman and Kandel (1987), de Roon and Nijman (2001), and Kan and Zhou (2012b). The mean-variance frontier can by definition only shift outwards when a set of assets is added to the investment universe. However, mean-variance spanning tests can be used to check whether a shift of the mean-variance frontier is too large to be attributed to chance. Performing statistical tests allows us to measure and compare diversification benefits from international bonds, stocks, and real estate by means of statistical significance.

Taking the perspective of a US investor, we start with a mixed-asset portfolio based on US bonds, US stocks, and US real estate. In three steps, we successively add international bonds, international stocks, and international real estate to this portfolio. We test the contribution of the different assets by two settings – first, by using currency-unhedged returns and, second, by using fully (unitary) currency-hedged returns – in order to account for additional systematic movements in currency markets.

If a specific currency loads on a common currency factor, this may introduce unintended co-movements between asset classes within a foreign country, measured in USD. If the currencies of several countries simultaneously load in the same direction on a common currency factor, unintended co-movements between investments in these countries might follow. Thus, if investors consider international assets on an unhedged basis to derive an asset allocation decision, the core asset price risk and the exchange rate risk are being taken as inseparable, and the "full diversification benefits of international investing" can be masked (see Dales and Meese (2001), p. 10).

To evaluate the impact of currency exposure in international portfolios, we use fully (unitary) hedged international asset returns to contrast our results based on unhedged returns. For the hedged returns, we use forward foreign exchange market data and account for hedging costs in this regard. The use of hedged returns allows to (almost) separate asset price risk from exchange rate risk in a realistic and quite applicable manner, since currency markets are highly liquid, and it is thus relatively easy to unwind an unintended exposure in the forward foreign exchange market.[67] Accordingly, our results based on hedged returns should provide a more accurate picture on the diversification benefits of core assets, i.e. international bonds, stocks, and real estate.[68]

[67]Indeed, fully hedged international returns are directly available for retail investors by "currency-hedged" ETFs which have become available since 2010 (see e.g. "WisdomTree Launches Currency-Hedged ETF", The Wall Street Journal, January 12, 2010, available on http://www.wsj.com).

[68]Of course, it is tempting to broaden the scope to optimal currency hedging schemes (see e.g. Glen and Jorion (1993), de Roon, Nijman, and Werker (2003), and Campbell, de Medeiros, and Viceira (2010)) or to consider currency as an asset class of its own. Such currency-overlays could improve the portfolio performance in addition

The findings of our analysis related to international bond and stock markets show that mean-variance efficiency is only weakly rejected for currency-unhedged returns while the intersection hypothesis for the tangency portfolio is not rejected at all, which is in line with the results from previous studies like Eun, Lai, de Roon, and Zhang (2010) and Glabadanidis (2009), among others. However, the results are stronger for international real estate, indicating that the efficient frontier is shifted upwards significantly when international real estate is added to an existing portfolio consisting of international bonds and stocks. This finding is even strengthened when the currency risk exposure is fully hedged. The results from an out-of-sample analysis confirm our findings. Therefore, the contribution of international real estate to an internationally diversified bond and stock portfolio is statistically significant and for an investor also economically meaningful. Finally, we study time trends in the diversification benefits discussed. We find that diversification benefits from international real estate are, in comparison to common stocks, large and relatively stable over time. However, during the recent financial crisis, also the diversification benefits from real estate turn statistically insignificant and economically meager. Interestingly, international bonds equalize the reduced diversification benefits nearly one to one.

This paper is structured as follows. Section 4.2 provides a brief literature review. Section 4.3 discusses the empirical methodology and Section 4.4 describes the data characteristics. In Section 4.5, we report our empirical results and discuss the implications of our findings. In Section 4.6, we apply out-of-sample tests including investment frictions in form of short selling constraints as a robustness check, and study the time trends of diversification benefits of all three asset classes. Section 4.7 offers a summary and provides concluding remarks.

4.2 Literature

Portfolio diversification can mainly be improved in two dimensions. Investors can seek additional asset classes, with the rationale that returns across them are not perfectly correlated.

to the investment in the underlying asset classes (bonds, stocks, real estate). For tractability and to distinguish from other (but important) research questions, this study focuses on analyzing the diversification benefits of traditional assets (and international real estate in particular).

Furthermore, international diversification is intended to reduce the risk of a portfolio, in the hope that returns across countries are not perfectly correlated (e.g. Grubel (1968) and Solnik (1974)). Unfortunately, as international capital markets have become more integrated over the last decades, international asset markets are considered to be increasingly correlated with each other. Indeed, such a tendency is well documented for international stock markets (e.g. Eun and Lee (2010a)). Driessen and Laeven (2007) apply mean-variance spanning tests for stock markets from developed and developing countries. They document larger diversification benefits for developing countries than for developed countries, and find decreasing diversification benefits over the sample period from 1985 to 2002. Kan and Zhou (2012b) apply mean-variance spanning tests to study the benefits of international diversification for a US investor and test the improvement of the tangency portfolio when seven currency risk-unhedged international stock markets of developed countries are added to the investment universe of US bonds and stocks. They find a statistically significant improvement for the tangency portfolio only for the period from 1970 to 1988. By contrast, the improvement is statistically insignificant for the most recent period from 1989 to 2007, implying only weak diversification benefits from international stocks (similar results are confirmed by Glabadanidis (2009)). Indeed, Errunza, Hogan, and Hung (1999) show that US investors can mimick international diversification gains from common stocks using investments in domestic (US) multinational corporations and industry indices.

Eichholtz (1996) argues that real estate returns could exhibit lower international correlation than common stocks due to the local nature of real estate markets. In consequence of the circumstance that investing in direct real estate in an international environment is quite difficult and expensive, an alternative considered is securitized real estate, i.e., exchange-traded real estate operating companies and REITs. For the most part, such companies' main focus lies on owning and letting property. Thus, they reflect the local real estate markets but are still highly liquid real estate investment vehicles with low transaction costs and therefore are well comparable to bonds and common stocks. Recent studies by Oikarinen, Hoesli, and Serrano (2011) and Yunus, Hansz, and Kennedy (2012), among others, confirm that, first, securitized real estate proxies for direct real estate and, second, leads the direct real estate market. For that reason, using listed real estate seems to be an appropriate vehicle in particular for global asset allocation.

In the following we confine ourselves to only mentioning those studies covering real estate, which are closest to the subject of ours; Worzala and Sirmans (2003) give a more comprehensive review of the relevant literature. Chen, Ho, Lu, and Wu (2005) apply mean-variance spanning tests in a national setup with US-REITs, and find significant diversification benefits towards US common stocks over the period from 1980 to 2002. Chiang and Lee (2007) follow them, covering the period from 1980 to 2004, but also include US bonds as benchmark assets. They consider both US direct real estate and US securitized real estate as an additional portfolio diversifier, and find a significant improvement of the investment opportunity set for direct real estate in mixed-asset portfolios. However, they are not able to make the same decisive conclusion for securitized real estate. Rubens, Louton, and Yobaccio (1998) measure diversification gains from international investments in Japan, the UK, and the US from 1978 to 1993. They cover US bonds, US stocks, and US direct real estate as benchmark assets. However, only bonds and stocks are considered international assets. The currency risk is unhedged and diversification benefits are reported insignificant for the tangency portfolio. Similar findings are reported by Glabadanidis (2009) and Kan and Zhou (2012b).

Several studies have examined factor models, exploring global or common risk factors driving returns of international securitized real estate. Bond, Karolyi, and Sanders (2003) study currency-unhedged real estate return characteristics from 14 countries over the period from 1990 to 2001, and identify country-specific factors, indicating potential diversification benefits. Liu and Mei (1998) also use a factor model to analyze securitized real estate and common stocks from Australia, France, Japan, South Africa, the UK, and the US for the period from 1980 to 1991. They compare currency-unhedged as well as hedged returns, and compute mean-variance efficient portfolios for both cases. Their results indicate diversification benefits for international real estate, but they do not provide evidence of statistical significance.

Even though there is a wide range of studies focusing on diversification benefits from international mixed-asset portfolios, there is a gap in the literature. The literature does not cover the statistical testing of the significance of diversification benefits from investments in international securitized real estate markets, while at the same time explicitly considering the impact of investors' exchange rate risk exposure. By applying mean-variance efficiency tests in the context of mixed-asset portfolios based on a broad range of national markets and three different assets, the contribution of this paper is to narrow this gap and to shed further light on this topic against the background of existing research.

4.3 Methodology

4.3.1 Spanning Tests

We apply the regression-based intersection and spanning tests proposed by Huberman and Kandel (1987) and Jobson and Korkie (1989) to measure diversification benefits from international real estate and other assets. If the mean-variance frontier constructed from some benchmark assets coincides with the frontier for a set of additional test assets, the benchmark assets span the frontier of all assets, and it is not possible to improve the investment opportunity set with the test assets. However, given a specific sample, the mean-variance frontier of the broader set of assets by definition shifts outwards with respect to the frontier of the smaller set. The following tests measure the statistical significance and indicate whether a shift of the frontier is too large to be attributed to chance. In accordance with the literature we distinguishe between testing "spanning", i.e. testing a shift of the *complete* mean-variance frontier, and testing "intersection", i.e. testing a shift of the mean-variance frontier at a pre-specified *single point* that is of special interest to the investor. In particular, we will consider the global minimum variance portfolio and the tangency portfolio as regions of special interest.

Consider a vector of $N + K$ asset returns, $1 + r_t = R_t = (R_{Kt}, R_{Nt})$, where R_{Kt} are K benchmark asset returns and R_{Nt} are N test asset returns at time t. Spanning and intersection tests can be based on the regression of the N test assets on the K benchmark assets,

$$R_{Nt} = \alpha + \beta R_{Kt} + \varepsilon_t. \tag{31}$$

Huberman and Kandel (1987) show that intersection can be tested by the N restrictions,

$$H_0: \ \alpha v + \beta 1_K - 1_N = 0_N, \tag{32}$$

for a given value of v. Setting v equal to the inverse of the risk-free rate, $v = 1/R_f$, is an intersection test of the tangency portfolio. If the coefficient restriction above cannot be rejected, the mean-variance frontier of the benchmark assets and the mean-variance frontier with the additional test assets intersect at the tangency portfolio and it is not possible to improve the portfolio performance at this point of the frontier. Similarly, setting v equal to zero is an intersection test at the global minimum variance portfolio in the mean-variance space. Spanning can be examined by testing intersection for all possible values of v, and is equivalent to the $2N$ restrictions $\alpha = 0_N$ and $\beta 1_K - 1_N = 0_N$.

If the disturbances ε_t are independent and identically distributed (i.i.d.), intersection ($\# = 1$) and spanning ($\# = 2$) can be tested with an exact F-test, $F \sim F_{\#N,\#(T-N-K)}$. When the distributions of the ε_t are non-i.i.d., the generalized method of moments (GMM) with a Newey-West covariance estimator can be used for a heteroscedasticity- and autocorrelation-robust (HAC) asymptotic Wald test, $W_{hac} \sim \chi^2_{\#N}$.[69] A survey on mean-variance efficiency test statistics can be found in de Roon and Nijman (2001). Implementation issues are discussed in detail by Kan and Zhou (2012b).

4.3.2 Currency Risk and Hedging

We apply mean-variance efficiency tests conditional on currency-unhedged and fully currency-hedged returns. Thereby, we will account for hedging costs investors face by using forward

[69]Bekaert and Urias (1996) propose a stochastic discount factor (SDF)-based test for mean-variance efficiency. They also show that the restrictions imposed by the SDF-based test are equivalent to the restrictions imposed by Huberman and Kandel (1987). The SDF-based tests and the regression-based tests allow us to draw almost identical conclusions, results are available from the authors on request.

exchange rate market data. This section is devoted to showing how we disentangle currency risk exposure from core asset price risk. For S_t being the spot US dollar price of one unit of a foreign currency, and P_t^* the price of the foreign asset inclusive of reinvestments in the local currency the currency-unhedged returns in US dollars are measured as

$$r_{t+1} = \frac{P_{t+1}^* S_{t+1}}{P_t^* S_t} - 1. \tag{33}$$

The unhedged US dollar return can be decomposed in a local currency asset return, $r_{t+1}^* = P_{t+1}^*/P_t^* - 1$, an exchange rate component, $s_{t+1} = S_{t+1}/S_t - 1$, and a cross product component $r_{t+1}^* s_{t+1}$, resulting in:

$$r_{t+1} = r_{t+1}^* + s_{t+1} + r_{t+1}^* s_{t+1}. \tag{34}$$

To hedge currency risk of a foreign asset, forward contracts can be used. Let $F_{t|t+1}$ be the one-period forward price of the exchange rate, then the return of a currency long-forward contract is $r_{F|t+1} = \left(S_{t+1} - F_{t|t+1}\right)/S_t$, and the forward premium is denoted as $f_{t|t+1} = F_{t|t+1}/S_t - 1$. In general, a foreign asset return given an arbitrary hedging strategy ψ is:

$$r_{t+1}^H = r_{t+1} - \psi r_{F|t+1}. \tag{35}$$

A simple hedging strategy is to unwind any passive currency holdings, i.e. $\psi = 1$. We use Equation (35) in the following to calculate fully (unitary) hedged returns of international assets. In this case, the currency-hedged return can be written as:

$$r_{t+1}^H = r_{t+1}^* + f_{t|t+1} + r_{t+1}^* s_{t+1}. \tag{36}$$

The difference between the unhedged return ($\psi = 0$) and the fully hedged return ($\psi = 1$) is that the uncertain exchange rate component of the return is being substituted by the certain forward premium. Thus, it is expected that the volatility of hedged returns is smaller than that of unhedged returns. In contrast, the impact on the sample mean depends on the sign of the average forward premium and can be positive as well as negative in general.

The unhedged as well as the fully hedged strategies are ad hoc currency management strategies. Jorion (1994), for instance, shows how an optimal currency strategy for a mean-variance investor can be analytically determined. Two components affect the optimal allocation with respect to currency risk exposure. The first is a speculative demand, determined by the risk-return ratio of the currency forward returns, and the second is a hedging demand, determined by the correlation structure to the core assets. Based on this, it is possible to show that an unhedged currency strategy ($\psi = 0$) is mean-variance optimal only, if currency forward returns have a zero expected return and are uncorrelated with the unhedged core assets. A comparison of Equation (33) and $r_{F|t+1}$ illustrates that this case is rather unrealistic, since both naturally contain the exchange rate component. An often proposed hedging strategy (Eun and Resnick (1988) and Perold and Schulman (1988)) is the full hedge ($\psi = 1$), and, as can be shown, it is also justified as an optimal hedge if the currency forward returns have a zero expected return and are uncorrelated with the core assets measured in the local currency.[70]

In this study, we are primarily interested in the core assets (i.e. international real estate), and we want to shut off potential effects from systematic foreign exchange movements (as documented by Verdelhan (2011)). For that reason, we focus on fully hedged returns and contrast them to the unhedged scenario. Studying fully hedged returns has the advantage that it is possible to dissect diversification benefits from the pure core asset component almost without the influence of ex ante currency risk.

[70]The empirical study by Campbell, de Medeiros, and Viceira (2010) finds that, generally, currency returns are often non-zero and correlated with some core assets measured in the local currency. For example, Campbell, de Medeiros, and Viceira (2010) find that an empirically mean-variance optimal hedge is close to a full hedge for international bonds. But there are some further gains for stock portfolios from overhedging positively correlated currencies and vice versa for negatively correlated currencies.

4.4 Data and Descriptive Statistics

The empirical analysis of spanning is based on the nine largest securitized real estate markets. Even though their market share has decreased during the last 20 years due to the fast growth of securitized real estate markets around the world, these markets still cover around 75% of global market capitalization in securitized real estate markets and large parts of international stock market capitalization as well as major government bond markets. Thus, market coverage is representative for US investors' investment universe. In addition to the US market, the markets in Australia, Canada, France, Hong Kong, Japan, the Netherlands, Singapore, Sweden, and the UK are covered and thus, spread through economic and geographic regions. The sample period ranges from 1984 to 2010, covering 27 years, which is an extensive time period compared to previous studies.

For the national securitized real estate markets we use monthly data from Global Property Research (GPR), while the national stock markets are represented by the MSCI country indices. National government bond markets are represented by the total return of bonds with a constant maturity of 10 years. We calculate total return bond indices using the yield-based approximation suggested by Campbell, Lo, and MacKinlay (1997) and Campbell, de Medeiros, and Viceira (2010). Bond yield data are from the International Financial Statistics database of the IMF. Except for the bond markets in Hong Kong and Singapore, all data are available for the whole sample period and thus, the three different asset indices for ten national markets build up a representative environment with 324 observations for each asset in each national market. Government bonds are not issued for large parts of the sample period for Singapore and there is no government bond market in Hong Kong. However, we consider these two real estate and stock markets in our analysis since both markets have a highly capitalized and well developed real estate sector. Referring to the considered securitized real estate indices, Serrano and Hoesli (2009) conclude that the GPR indices are well suited to both measure the performance of the market and evaluate portfolio performance.

The monthly returns from the total return indices are calculated as simple discrete returns.

The average monthly returns and corresponding standard deviations in Table 28 are reported in US dollars for unhedged and fully currency-hedged assets. Fully hedged returns are computed according to Equation (35) using foreign spot and forward exchange rates from Barclays Bank and WM Reuters available on Datastream.

Considering average returns for the US market, securitized real estate performs better than government bonds and slightly better than stocks, while government bonds show a much lower volatility than both stocks and real estate. Across national asset markets, Australia, France, Hong Kong, Singapore, and Sweden show a strong performance for all assets over the sample period, while bonds from the UK, Japanese stocks, and Canadian real estate have the lowest average returns. Compared to the US market, the performance of the international government bond and stock markets seems to be stronger at first glance. However, the US dollar has depreciated against many currencies in the last 27 years and thus, some of the return of the unhedged indices is attributed to currency gains. This finding is also confirmed when comparing unhedged and fully hedged returns. For all covered markets, fully currency-hedged returns are lower than unhedged returns. At the same time, however, volatility also decreases substantially and is mostly comparable to the volatility of the corresponding US asset market. Thus, naïve currency hedging over the sample period results in lower returns and substantially lower volatility. Exceptions are, again, Hong Kong and Singapore. While the Hong Kong dollar is fixed against the US dollar, the Singapore dollar is freely traded but pegged by a basket of other currencies. Consequently, the statistics also show that return volatility does not substantially decrease for Hong Kong and Singapore when fully currency-hedged returns are considered. Related to the bond markets, their strong risk-adjusted performance is mainly driven by decreasing interest rates over a wide range of the sample period and particularly in the aftermath of the financial market turmoil at the end of the sample period.

Table 28: Descriptive Statistics of Monthly Total Returns

The table reports the mean and standard deviation (StD) for US and international bonds, stocks and securitized real estate in percent. Below, we report the condition index of parts and for the complete correlation matrix of asset returns. The condition index is defined as $(maximum\,root/minimum\,root)^{1/2}$. The returns from a fully currency-hedged strategy are calculated as described in Section 4.3.2. Returns are in USD. The sample period ranges from 01/1984 to 12/2010.

	Bonds (10y)				**Stocks**				**Real Estate**			
	Mean	StD	Mean	StD	Mean	StD	Mean	StD	Mean	StD	Mean	StD
National Assets												
US	0.72	1.95			0.95	4.53			1.03	5.59		
International Assets												
	unhedged		fully hedged		unhedged		fully hedged		unhedged		fully hedged	
Australia	0.97	4.08	0.61	2.07	1.21	6.76	0.85	4.86	1.13	5.87	0.77	4.27
Canada	0.88	2.81	0.73	1.87	1.00	5.66	0.84	4.56	0.31	7.14	0.15	6.33
France	0.97	3.53	0.72	2.31	1.20	6.29	0.95	5.97	1.20	5.85	0.94	4.92
Japan	0.98	4.59	0.84	2.53	0.71	6.59	0.57	5.87	1.21	10.22	1.07	9.51
Netherlands	0.86	3.59	0.67	2.13	1.20	5.68	1.01	5.54	0.88	5.18	0.70	4.45
Sweden	0.93	3.69	0.68	2.00	1.44	7.53	1.19	7.01	1.28	9.77	1.04	9.58
UK	0.83	3.78	0.59	1.93	1.02	5.32	0.78	4.73	0.88	6.78	0.63	6.03
Hong Kong					1.59	8.08	1.61	8.06	1.97	10.81	1.99	10.79
Singapore					0.95	7.62	0.92	7.07	1.58	11.52	1.55	10.94
Correlation Matrix												
	unhedged		fully hedged		unhedged		fully hedged		unhedged		fully hedged	
Cond. index	9.98		6.72		5.95		6.00		4.86		3.91	
	all assets, unhedged: 17.79						all assets, fully hedged: 14.10					

When comparing returns and volatility from international real estate with those from the US securitized real estate market, international markets do not seem to provide substantial diversification benefits at first glance, but correlations are completely neglected at this stage.

In addition to the risk and return characteristics, correlations also differ between the three asset classes. To summarize the correlation matrix in one number (and to save space), we also compute the condition index of the correlation matrix in Table 28.[71] The condition index is defined as the $(maximum\, root/minimum\, root)^{1/2}$, and is traditionally used to detect multicollinearity. In principle, a larger condition index indicates more correlated data and a lower condition index indicates less correlated data.

Comparing the three asset classes, we find that the international real estate markets have a lower condition index than international stocks and bonds. This finding suggests that international diversification could work better for real estate than for stocks, as proposed by Eichholtz (1996). We also compare the condition index of the three asset classes for unhedged and fully hedged returns. For international bonds and international real estate, the condition index is substantially lower when the currency risk is fully hedged. In contrast, based on the condition index, we cannot find an improvement for international stocks when fully hedged returns are compared to unhedged returns, as the condition index remains almost unchanged. We provide the condition index based on all assets for both hedging schemes to trace out the effect of currency hedging on the entire correlation matrix of the considered assets. Again, we find a substantial drop of the condition index from 18 for unhedged assets to 14 for fully hedged assets.

In summary, our findings seem to be reasonable. Hedging removes currency specific risk present in any dollar/foreign currency pair and therefore, correlations between currency-unhedged returns contain an additional co-movement which is not caused by the core assets themselves. These findings square well with Verdelhan (2011), who documents substantial systematic variation in bilateral exchange rates. Accordingly, we expect more remarkable diversification benefits for fully hedged international returns than for unhedged returns.

4.5 Empirical Results

Domestic Diversification. The empirical results from regression-based spanning and intersection tests are presented in Table 29. Additionally, the maximum Sharpe ratio and the increase of the Sharpe ratio with respect to the benchmark assets are reported for each setting. First, in

[71] We thank Paskalis Glabadanidis for this suggestion.

Panel A, we consider a US investor who holds US bonds and US stocks, and considers expanding the investment universe with US securitized real estate. According to the F-test as well as the W_{hac}-test, the null hypothesis of spanning cannot be rejected at any common significance level. This means that the investor cannot significantly improve her portfolio with US securitized real estate. The spanning tests are confirmed by the intersection tests for the global minimum variance portfolio as well as for the tangency portfolio. Accordingly, in economic terms, we find an increase in the Sharpe ratio close to zero.

International Diversification with Unhedged Currency Risk. Panel B of Table 29 reports the results from mean-variance spanning tests and intersection tests based on currency-unhedged returns of international assets. This means that in addition to bond, stock, and real estate market risk, currency risk exposure is still present in the portfolios. To be as strict and conservative as possible with the test results related to real estate, which is one of our main topics in the analysis, we first conduct mean-variance efficiency tests for bond and stock markets, before expanding the investment universe with international real estate. From a practical point of view, this procedure is not implausible, since most investors may add international bonds and stocks to their domestic mixed-asset portfolio before they consider investing in international real estate. Therefore, while spanning for international bonds is based on three benchmark assets (US bonds, US stocks, and US real estate), spanning on international real estate is based on a challenging benchmark of 19 assets (three US assets, seven international bond markets, and nine international stock markets).

Table 29: Results from Mean-Variance Spanning and Intersection Tests

The table reports p-values of spanning test statistics for a complete shift of the mean-variance frontier when N test assets are added to K benchmark assets. We also test for intersection, i.e., a shift at a single point of the mean-variance frontier, where TP is the tangency portfolio and GMVP is the global minimum variance portfolio. We proxy the risk-free rate by the US T-bill rate averaged over the sample period (0.38% per month). SR is the Sharpe ratio when the test assets are included, and $\triangle SR$ is the corresponding differential Sharpe ratio obtained by adding the test assets to the benchmark assets. The sample period ranges from 01/1984 to 12/2010.

	Spanning		**Intersection GMVP**		**Intersection TP**			
	F	W_{hac}	F	W_{hac}	F	W_{hac}	SR	$\triangle SR$
Panel A: domestic diversification								
benchmark portfolio: US bonds & US stocks (K=2)								
US real estate (N=1)	0.274	0.334	0.137	0.218	0.763	0.776	0.21	0.00
Panel B: international diversification - currency risk unhedged								
benchmark portfolio: US assets (K=3)								
int. bonds (K=3, N=7)	0.025	0.009	0.005	0.003	0.369	0.270	0.26	0.05
+ int. stocks (K=10, N=9)	0.201	0.249	0.055	0.136	0.642	0.504	0.30	0.04
+ int. real estate (K=19, N=9)	0.003	0.001	0.002	0.000	0.098	0.098	0.38	0.08
all int. assets (K=3, N=25)	0.001	0.000	0.000	0.000	0.224	0.102	0.38	0.17
Panel C: international diversification - currency risk fully hedged								
benchmark portfolio: US assets (K=3)								
int. bonds (K=3, N=7)	0.000	0.000	0.000	0.000	0.242	0.159	0.27	0.06
+ int. stocks (K=10, N=9)	0.001	0.000	0.000	0.000	0.642	0.498	0.31	0.04
+ int. real estate (K=19, N=9)	0.000	0.000	0.000	0.000	0.061	0.058	0.39	0.08
all int. assets (K=3, N=25)	0.000	0.000	0.000	0.000	0.127	0.079	0.39	0.18

Spanning for international bonds is rejected at the 5% level for the F-test, and the W_{hac}-test. Considering specific points of the mean-variance frontier with intersection tests, the null hypothesis is rejected for the global minimum variance portfolio, but not for the tangency portfolio at the 5% level. The increase of the monthly Sharpe ratio, reported in the last column, is about 0.05 and indicates that the shift of the tangency portfolio is large in economic terms, at least.

In the next row, we add international stocks to the investment menu. It is not possible to reject spanning at common levels for the F-test and the W_{hac}-test. A closer look at the global minimum variance portfolio and the tangency portfolio shows that diversification benefits seem to be low especially in the region close to the tangency portfolio with p-values of 0.64 and 0.50. However, the increase of the monthly Sharpe ratio is about 0.04 and again respectable in economic terms. Our results are in line with the findings by Kan and Zhou (2012b) for the period from 1989 to 2007 (which is the closest to our sample period) and Glabadanidis (2009), who also find large but insignificant changes in the tangency portfolio.

By contrast, the empirical results from the spanning tests and intersection tests are stronger for international (securitized) real estate than for international stocks. This is confirmed by p-values below 0.01 for the spanning tests as well as for the intersection tests for the global minimum variance portfolio. The intersection tests for the tangency portfolio have p-values close to but below 0.10. Thus, the increase of the Sharpe ratio of 0.08 is also significant for international real estate at the 10% level.

In the last row of Panel B, we test all 25 international assets against the benchmark of three US assets. The rejection of spanning provides evidence that international diversification yields substantial benefits. However, as the intersection tests indicate, the diversification benefits from international investing are only confirmed for the global minimum variance portfolio. When turning to the tangency portfolio we find that an almost doubling of the Sharpe ratio by investing internationally is however not statistically significant at the 10% level. Hence, our results provoke the conclusion that the well-known home bias in asset allocation decisions is rational for an investor who is interested in the tangency portfolio and considers currency-unhedged returns for investment decisions.

Summarizing the results from the several spanning and intersection tests for currency-unhedged returns, it can be stated that spanning is not rejected for international stocks. We find significant diversification benefits from international bonds and international real estate at the region close to the global minimum variance portfolio. For the economically interesting tangency portfolio, only international real estate provides at least weakly significant diversification benefits.

However, at this moment one substantial risk factor in international portfolio diversification is not considered, namely the exposure to currency risk. Therefore, the offered benefits from a diversification across assets and across national markets may be misleading and the contribution of distinct assets may be biased by systematic movements in exchange rates generating additional correlation, volatility and/or returns.

International Diversification with Fully Hedged Currency Risk. Panel C of Table 29 provides mean-variance efficiency tests in the same order as in Panel B, but with fully currency-hedged international returns. The differences for international diversification in Panel C are overall remarkable. Based on the benchmark of three US assets, all test statistics strongly reject spanning and intersection at the global minimum variance portfolio for international bonds, international stocks and international real estate at the 1% level.

When we turn to the tangency portfolio, we find an additional improvement of the Sharpe ratio in each setting compared to the corresponding setting with unhedged currency risk. We find considerable lower p-values for the intersection tests of the tangency portfolio for international bonds and for international real estate. However, we can still not reject intersection for international bonds and international stocks. This finding conjectures that both bond markets and stock markets are exposed to common risk factors. These risk factors may not differ substantially for international bonds and US assets on the one side and international stocks on the other.

Against this background, the statistical results from the spanning tests for international real estate are even more remarkable and strongly support the conjecture that investments in international real estate also yield significant diversification benefits for the tangency portfolio. The shift of the tangency portfolio is statistically significant with p-values of 0.06 for both test statistics and the increase of the Sharpe ratio from 0.31 to 0.39 is also economically meaningful.

Summarizing the results above, we find that investors are well advised to, first, consider adding international real estate to their asset allocation and to, second, pay attention to currency exposure. After accounting for currency risk, the improvement of the tangency portfolio is not only economically huge, measured by the Sharpe ratio, but is also statistically significant, and leads to the conclusion that a home bias in asset allocation decisions is costly, in contrast to the results from unhedged returns. Furthermore, and related to the systematically higher Sharpe ratios from fully currency-hedged returns, it may be possible that an optimal currency hedging strategy yields further diversification benefits (see Campbell, de Medeiros, and Viceira (2010)).

Graphical Interpretation. Figure 12 presents a graphical illustration of the results discussed above for the fully currency-hedged portfolios. It sheds further light on our findings in the traditional representation of the mean-variance framework. Considering single asset markets, the group of the seven bond markets is relatively homogeneous in their risk-return profile, compared to the much more heterogeneous stock markets and real estate markets. The outward shift of the efficient frontier resulting from the addition of international bonds to a US mixed-asset portfolio is distinguishable, while there is only a modest outward shift in the mean-variance frontier when international stocks are included. This result is in line with the findings from spanning tests which are less significant, if at all, compared to bonds. Finally, there is a notable outward shift of the mean-variance efficient frontier when international real estate markets are taken into consideration for optimal portfolio allocation.

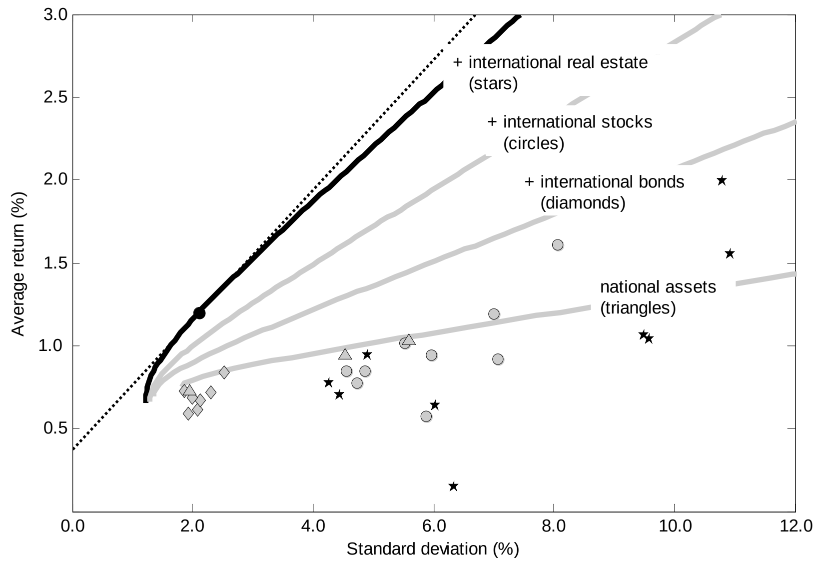
4.6 Further Results

Out-of-Sample Sharpe Ratios and Short Selling Constraints. In this section, we ask whether an investor can also achieve the reported diversification benefits in real time and take account of possible investment restrictions (short selling constraints). As shown in the previous section, improving the Sharpe ratio is the most challenging task and is thus likely to be also the most rigorous robustness test. We narrow our analysis on international securitized real estate due to the focus of the paper.

We use the first 180 observations of our sample to calculate mean-variance portfolio weights for the tangency portfolio. We hold this portfolio for one period, collect the return and move the rolling window one month forward to obtain new optimal portfolio weights, and so forth.

Figure 12: Mean-Variance Frontiers for Fully Currency-Hedged Returns

The figure plots the mean-variance efficient frontiers for US assets (triangles) and frontiers successively augmented with fully hedged international bonds (diamonds), international stocks (circles), and international securitized real estate (stars). The sample covers monthly data, the period ranges from 01/1984 to 12/2010.



We do this at first considering only the benchmark assets and subsequently the benchmark assets together with the test assets. In particular, the benchmark consists of the three US assets (US bonds, US stocks, US real estate), international bonds, and international stocks, and the augmented portfolio adds international real estate to the investment universe. From this procedure we obtain a time-series of 144 out-of-sample returns with and without international real estate.[72]

As the portfolio formation rule has substantial influence on the out-of-sample results (see e.g. DeMiguel, Garlappi, and Uppal (2009)), it is of particular importance to consider multiple rules to draw a picture of the feasible diversification benefits from some test assets. Therefore, we use several portfolio formation rules as they are described in detail by DeMiguel, Garlappi, and Uppal (2009). We consider the classical mean-variance tangency portfolio, the classical mean-variance tangency portfolio with a Bayes-Stein shrinkage estimator, the mean-variance tangency

[72]Similar out-of-sample tests are also conducted by de Roon, Nijman, and Werker (2003) and Eun, Lai, de Roon, and Zhang (2010).

portfolio with short-selling constraints[73], the mean-variance portfolio ignoring average returns (i.e. the global minimum variance portfolio), and a simple portfolio with equal weights ("1/N") for all assets.

We report out-of-sample Sharpe ratios of unhedged and fully hedged international assets in Table 30. We find an improvement of the Sharpe ratio due to international real estate in each setting, i.e. the investor was never worse off. Such a positive finding is not self-evident. In an out-of-sample setting, and in contrast to in-sample tests, the Sharpe ratio of the augmented portfolio could worsen with respect to the benchmark. The increase of the out-of-sample Sharpe ratio when international real estate is included is also economically meaningful, ranging from an additional 7% return per unit of risk for the short-sales constrained mean-variance tangency portfolio to up to 30% for the portfolio using the global minimum-variance portfolio weights. Across all portfolio formation rules, the latter also leads to the highest out-of-sample Sharpe ratio of 0.35, when international assets are fully hedged and international real estate is added.

Below the Sharpe ratio of each augmented portfolio which includes international real estate, we also provide HAC-robust t-statistics of the difference of the Sharpe ratio to the benchmark portfolio in brackets, proposed by Ledoit and Wolf (2008). We find statistically significant improvements for the portfolio based on the simple mean-variance tangency, the Bayes-Stein, and the global minimum-variance rule (10% level or lower). Again, we find the largest t-statistic for the latter of about 2.9 (2.5) using fully hedged (unhedged) international assets. Finally, the effect of currency hedging is also confirmed out-of-sample. Throughout the table, we find higher Sharpe ratios for the fully hedged than for the equivalent unhedged allocations.

[73]According to Bris, Goetzmann, and Zhu (2007), short selling is in principle allowed and practiced in all ten countries covered in our analysis for the major time span of our sample period. Nevertheless, in this section we include market frictions and/or investment restrictions (e.g. for pension funds), in form of short selling constraints to the mean-variance efficiency tests, as a further robustness check.

Table 30: Out-of-Sample Sharpe Ratios

The table reports out-of-sample Sharpe ratios for a “Benchmark” and a “+IRE” allocation. The “Benchmark” allocation consists of US bonds, US stocks, US real estate, international bonds, and international stocks. The “+IRE” allocation adds international real estate to the assets covered in the “Benchmark” group. We use the first 180-month of our sample to calculate optimal portfolio weights. We hold this portfolio for one period to obtain the first out-of-sample return. Next, we move the rolling window one period forward and repeat the procedure. This leads to 144 out-of-sample returns in total, which we use to calculate the out-of-sample Sharpe ratio. We proxy the risk-free rate by the US T-bill rate averaged over the sample period. We report HAC robust t-statistics for the difference Sharpe ratio between the “+IRE” and the “Benchmark” allocation in brackets. The sample period ranges from 01/1984 to 12/2010.

	unhedged		**fully hedged**	
	Benchmark	**+IRE**	**Benchmark**	**+IRE**
	Out-of-sample Sharpe ratios			
Equal weights “1/N”	0.19	0.21	0.22	0.24
		[0.99]		[0.90]
Mean-variance TP	0.08	0.13	0.18	0.24
		[1.30]		[1.59]
Mean-variance TP - Bayes-Stein	0.16	0.21	0.23	0.30
		[1.57]		[2.06]
Mean-variance TP - short constr.	0.19	0.20	0.28	0.30
		[0.86]		[0.78]
Mean-variance GMVP	0.23	0.30	0.27	0.35
		[2.48]		[2.94]

Time Trends in the Diversification Benefits. We examine in this section if the previously found diversification benefits from international assets, especially from international real estate, contain any time trend. It is well documented in the literature that linkages between assets and national markets are time-varying, that financial markets have become more integrated in the last decades, and that, as a consequence, international diversification benefits are possibly decreasing (e.g. Ang and Bekaert (2002), Bekaert and Harvey (1995), Bekaert and Harvey (2005), Goetzmann, Li, and Rouwenhorst (2005), Longin and Solnik (1995), and Longin and Solnik (2001)). Concerning to real estate in a mixed-asset context, Sa-Aadu, Shilling, and Tiwari (2010) find that real estate – similar to precious metals – provides good hedging characteristics against adverse shocks to consumption growth opportunities and when the economy is in a bad state.

To carry out our analysis, we follow the methodology of Eun, Lai, de Roon, and Zhang (2010) and focus on the tangency portfolio. We construct forward-rolling subsamples of 180 months, proceeding in one-month intervals. The first of 144 subsamples spans the time period from 01/1984 to 12/1998, the last subsample the period from 01/1996 to 12/2010. For each subsample, we follow the same order as in Table 29. First, we add international bonds to a benchmark of three national assets (US bonds, US stocks, US real estate), and collect the differential Sharpe ratio. Subsequently, we treat international bonds and national assets as benchmark assets, and add international stocks as test assets. In the same way, we next add international (securitized) real estate as test assets, and finally we treat all three international asset classes combined as test assets against the benchmark containing only national assets. From this procedure, we obtain a time series of differential Sharpe ratios of international assets against national assets, and can control for which part of diversification benefits can be attributed to which asset class over time.

The contribution to the differential Sharpe ratio of the augmented portfolio with international assets is visualized over time in Figure 13 and Table 31 presents the according statistics. Our discussion focuses on the results from *fully currency-hedged* returns. As can be seen, international bonds deliver steadily rising diversification benefits, with a highly significant time trend. This is not surprising, since long-term bond prices have been rising in the past decades due to falling long-term yields. The, in historical terms, exceptional risk-adjusted performance of long-term bonds over the past years has already been discussed in the data section of the paper (e.g. Palazzo and Nobili (2010) for a further discussion). However, Figure 13 reveals that diversification benefits from long-term bonds can be quite low in an economic environment with globally high or rising long-term yields, as in the late 1980s or early 1990s, covered mainly in the first 60 subperiods. The comparison to the unhedged statistics also shows that currency hedging has a rather dramatic positive impact on the diversification benefits from international bonds.

Table 31: Statistics for Time Trends in the Differential Sharpe Ratio

The table reports descriptive statistics and time trend regressions of differential Sharpe ratios (ΔSR) of the tangency portfolio between a benchmark portfolio of national (US) assets and the augmented portfolio. For each set, we compute the Sharpe ratio for 180-month windows, rolling forward one-month at a time, resulting in 144 observations (obs) obtained from the sample period from 01/1984 to 12/2010. Sign refers to the number of significant ΔSR at the 10% level under the W_{hac} test statistic for intersection at the tangency portfolio. We report t-statistics in brackets that are robust against heteroscedasticity and serial autocorrelation (Newey-West, three lags).

Descriptive statistics of $\triangle SR_t$						
	Mean	StD	sign/obs	Mean	StD	sign/obs
	unhedged			fully hedged		
intern. bonds	0.05	0.02	0/144	0.10	0.04	62/144
+ intern. stocks	0.09	0.02	39/144	0.07	0.02	12/144
+ intern. real estate	0.14	0.03	113/144	0.14	0.04	113/144
all intern. assets	0.28	0.02	112/144	0.31	0.02	136/144

Time trend regressions: $\triangle SR_t = \alpha + \beta\, time_t + \varepsilon_t$						
	α	β	adj. R^2	α	β	adj. R^2
	unhedged			fully hedged		
intern. bonds	0.0481	0.0000	0.00	0.0465	0.0008	0.65
	[7.38]	[0.44]		[4.84]	[6.32]	
+ intern. stocks	0.0981	-0.0001	0.03	0.0753	-0.0001	0.04
	[13.85]	[-0.97]		[11.91]	[-1.08]	
+ intern. real estate	0.1424	-0.0001	0.01	0.1794	-0.0005	0.28
	[17.36]	[-0.66]		[14.93]	[-2.82]	
all intern. assets	0.2887	-0.0002	0.09	0.3011	0.0002	0.12
	[45.32]	[-1.69]		[65.10]	[2.70]	

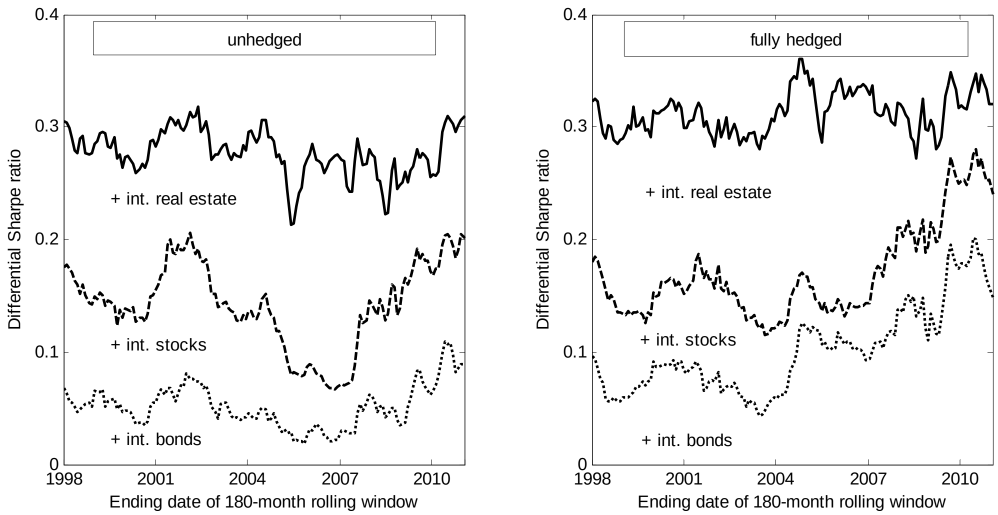
The diversification benefits measured in economic terms by differential Sharpe ratios from international stocks (visualized as the area between the two dotted lines in Figure 13) seem to be oscillating. The differential Sharpe ratios are rather large in subsamples with mainly bullish markets (mid 1990s, mid 2000s) and decrease in subsamples with mainly bearish markets (late 1980s, during the aftermath of the dotcom bubble burst in 2002, and most distinctively in the aftermath of the most recent global financial market turmoil). The time trend regressions indicate that there is a small and insignificant time trend of falling differential Sharpe ratios for international stocks.

Similarly, we find oscillating diversification benefits from international real estate, which are measured in *addition* to the benefits from international stock markets. The differential Sharpe ratios of international real estate are, on average, twice as large as those from international stock markets, which can also easily be seen in the figure (area between the continuous and the first dotted line). However, the additional diversification benefits decrease strongly in the subsamples ending after 2008. In line with this, we find that all 136 significant differential Sharpe ratios of the 144 subsamples end before 2008. The strong decrease of diversification benefits at the end of the time series also explains the significant negative time trend provided by the regression in Table 31. We see this as evidence that international real estate provides substantial diversification benefits in good times, and less (but still significant) benefits in bad times (e.g. in 2001). However, in very bad times, during the financial turmoil of 2008, international securitized real estate is not able to provide significant diversification benefits which are beyond that of international stocks. Clearly, international bonds can contribute most to the portfolio performance in such very bad times, as is shown at the end of the sample period in Figure 13. However, this period is also characterized by substantial interest rate decreases by central banks all over the world. Therefore, it is difficult to judge how representative this finding is for other periods, i.e., hot periods which are not accompanied by decreasing interest rates.

Finally, measuring international diversification in terms of all three asset classes together, we find a significant and positive time trend when the currency risk is fully hedged and potential systematic movements in exchange rates are shut off. Diversification benefits seem to be steadily swinging, whereas in some periods one or another asset class contributes more or less to the differential Sharpe ratio against a national portfolio. This does not hold regarding the currency risk-unhedged counterpart. In this setting, we find a significant decreasing time trend which may lead to the conclusion that diversification benefits from international investing decrease over time.

Figure 13: Time Trends in the Differential Sharpe Ratio

The figure plots time trends (based on 180-month windows, rolling forward one-month at a time) of differential Sharpe ratios. The benchmark allocation consists of US assets (US bonds, US stocks, and US real estate). The benchmark assets are successively augmented with international bonds, international stocks, and international securitized real estate. International assets are unhedged on the left side of the figure and they are fully hedged on the right side. The sample period ranges from 01/1984 to 12/2010.



4.7 Conclusion

In this paper, we investigate investors' benefits from international diversification for a sample of ten countries in a mixed-asset portfolio context including international bonds, stocks, and international securitized real estate, in particular from the perspective of a US investor. Furthermore, we have considered the impact of currency risk exposure, which is often neglected in the analysis of international diversification benefits, but could substantially influence the results since exchange rates exhibit systematic variation.

We find significant diversification benefits for global bond and stock portfolios from investing in international real estate. These benefits are robust in an out-of-sample analysis and available for most subperiods covered by our sample. The only exception are the subperiods covering the recent financial crisis, during which there are no diversification benefits achievable from international real estate investments. Overall, diversification benefits seem to be larger from international real estate than from common stocks. These findings support the view that real estate returns exhibit lower international exposure than common stocks and are less globally integrated due to the local nature of real estate markets (see Eichholtz (1996)). We also document that accounting for currency risk exposure of international assets is helpful to separate out the true diversification benefits from global bonds, stocks, and real estate.

Since the conducted analysis is based on historical data and applies a static or myopic framework, further research could contribute to answering the question of which variables are good predictors for investment decisions and whether the diversification benefits are still significant in a dynamic, time-varying framework. Regarding hedging currency risk and the presented evidence that currency hedging can matter for portfolio allocations, an analysis on optimal currency hedging should provide interesting insight into this topic. Against the background of the recent financial crisis in particular, constructing and conducting spanning tests using conditioning variables in an international context also presents an interesting topic for future research (see e.g. Ferson and Siegel (2010)).

References

ADRIAN, T., E. ETULA, AND T. MUIR (2011): "Financial Intermediaries and the Cross-Section of Asset Returns," Staff Report no. 464, Revised April 2011.

AIT-SAHALIA, Y., J. A. PARKER, AND M. YOGO (2004): "Luxury Goods and the Equity Premium," *Journal of Finance*, 59, 2959–3004.

ANDERSON, R. W., AND J.-P. DANTHINE (1981): "Cross Hedging," *Journal of Political Economy*, 89, 1182–1192.

ANDREWS, D. W. K. (1991): "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59, 817–858.

ANG, A., AND G. BEKAERT (2002): "International Asset Allocation with Regime Shifts," *Review of Financial Studies*, 15, 1137–1187.

ANG, A., AND J. S. CHEN (2010): "Yield Curve Predictors of Foreign Exchange Returns," Working Paper.

ANG, A., R. HODRICK, Y. XING, AND X. ZHANG (2006): "The Cross-Section of Volatility and Expected Returns," *Journal of Finance*, 61, 259–299.

ASNESS, C. S., J. M. LIEW, AND R. L. STEVENS (2000): "Parallels Between the Cross-Sectional Predictability of Stock and Country Returns," *Journal of Portfolio Management*, 23, 79–87.

ASNESS, C. S., T. J. MOSKOWITZ, AND L. H. PEDERSEN (2012): "Value and Momentum Everywhere," *Journal of Finance*, forthcoming.

BACCHETTA, P., AND E. VAN WINCOOP (2010): "Infrequent Portfolio Decisions: A Solution to the Forward Discount Puzzle," *American Economic Review*, 100, 870–904.

Bakshi, G., N. Kapadia, and D. Madan (2003): "Stock Return Characterisitics, Skew Laws, and the Differential Pricing of Individual Equity Options," *Review of Financial Studies*, 13, 101–143.

Balduzzi, P., and C. Robotti (2008): "Mimicking Portfolios, Economic Risk Premia, and Tests of Multi-Beta Models," *Journal of Business & Economic Statistics*, 26, 354–368.

Bansal, R., and A. Yaron (2004): "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *Journal of Finance*, 59, 1481–1509.

Barroso, P., and P. Santa-Clara (2012): "Beyond the Carry Trade: Optimal Currency Portfolios," Working Paper.

BEA (2009): *Concepts and Methods of the U.S. National Income and Product Accounts*. The NIPA Handbook, 2009.

Bekaert, G., and C. Harvey (1995): "Time-Varying World Market Integration," *Journal of Finance*, 50, 403–444.

——— (2005): "Market Integration and Contagion," *Journal of Business*, 78, 39–69.

Bekaert, G., and M. S. Urias (1996): "Diversification, Integration and Emerging Market Closed-End Funds," *Journal of Finance*, 51, 835–869.

Bell, W. R., and D. W. Wilcox (1993): "The Effect of Sampling Error on the Time Series Behavior of Consumption Data," *Journal of Econometrics*, 55, 235–265.

Belo, F. (2010): "Production-based Measures of Risk for Asset Pricing," *Journal of Monetary Economics*, 57, 146–163.

Bhojraj, S., and B. Swaminathan (2006): "Macromomentum: Returns Predictability in International Equity Indices," *The Journal of Business*, 79, 429–451.

BIS (2010): "Triennial Central Bank Survey," *Foreign Exchange and Derivative Market Activitiy in April 2010*, Revised September 2010.

Bond, S. A., G. A. Karolyi, and A. B. Sanders (2003): "International Real Estate Returns: A Multifactor, Multicountry Approach," *Real Estate Economics*, 31, 481–500.

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

Breeden, D. T. (1979): "An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities," *Journal of Financial Economics*, 7, 265–296.

Breeden, D. T., M. Gibbons, and R. Litzenberger (1989): "Empirical Tests of the Consumption-oriented CAPM," *Journal of Finance*, 44, 231–262.

Bris, A., W. N. Goetzmann, and N. Zhu (2007): "Efficiency and the Bear: Short Sales and Markets Around the World," *Journal of Finance*, 62, 1029–1080.

Britten-Jones, M. (1999): "The Sampling Error in Estimates of Mean-Variance Efficient Portfolio Weights," *Journal of Finance*, 54, 655–671.

Brunnermeier, M. K., S. Nagel, and L. H. Pedersen (2009): "Carry Trades and Currency Crashes," *NBER Macroeconomics Annual 2008*, 23, 313–347.

Burnside, C. (2010): "Identification and Inference in Linear Stochastic Discount Factor Models," NBER Working Paper 16634.

——— (2011): "The Forward Rate Premium is Still a Puzzle, a Comment on "The Cross-Section of Foreign Currency Risk Premia and Consumption Growth Risk"," *Amercian Economic Review*, 101, 3456–3476.

Burnside, C., M. Eichenbaum, and S. Rebelo (2011): "Carry Trade and Momentum in Currency Markets," *Annual Review of Financial Economics*, 3, 511–535.

Burnside, C. A., B. Han, D. Hirshleifer, and T. Y. Wang (2011): "Investor Overconfidence and the Forward Premium Puzzle," *Review of Economic Studies*, 78, 523–58.

Campbell, J. Y. (1996): "Understanding Risk and Return," *Journal of Political Economy*, 104, 298–345.

——— (2000): "Asset Pricing at the Millennium," *Journal of Finance*, 55, 1515–1567.

——— (2003): "Consumption-based Asset Pricing. In: Constantinides, G., Harris, M., Stulz, R. (Eds)," Handbook of the Economics of Finance, North-Holland, Amsterdam, 803-887.

Campbell, J. Y., K. S. de Medeiros, and L. M. Viceira (2010): "Global Currency Hedging," *Journal of Finance*, 65, 87–121.

Campbell, J. Y., A. W. Lo, and A. C. MacKinlay (1997): "The Econometrics of Financial Markets," Princeton University Press, New Jersey.

Campbell, J. Y., A. Sunderam, and L. Viceira (2010): "Inflation Bets or Deflation Hedges? The Changing Risks of Nominal Bonds," *Working Paper*.

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

Chen, H.-C., K.-Y. Ho, C. Lu, and C.-H. Wu (2005): "Real Estate Investment Trusts," *Journal of Portfolio Management*, Special Issue, 46–54.

Chen, N.-F., R. Roll, and S. A. Ross (1986): "Econmomic Forces and the Stock Market," *Journal of Business*, 59, 383–403.

Chiang, K. C. H., and M.-L. Lee (2007): "Spanning Tests on Public and Private Real Estate," *Journal of Portfolio Management*, 13, 7–15.

Chordia, T., and L. Shivakumar (2002): "Momentum, Business Cycle, and Time-Varying Expected Returns," *Journal of Finance*, 62, 985–1019.

Christiansen, C., A. Ranaldo, and P. Söderlind (2010): "The Time-Varying Systematic Risk of Carry Trade Strategies," *Journal of Financial and Quantitative Analysis*, 46, 1107–1125.

Christoffersen, P., V. Errunza, K. Jacobs, and X. Jin (2012): "Is the Potential for International Diversification Disappearing? A Dynamic Copula Approach," *Review of Financial Studies*, forthcoming.

Cochrane, J. H. (1991): "Production-based Asset Pricing and the Link Between Stock Returns and Economic Fluctuations," *Journal of Finance*, 46, 209–237.

——— (1996): "A Cross-Sectional Test of an Investment-Based Asset Pricing Model," *Journal of Political Economy*, 104, 572–621.

——— (2005): "Asset Pricing," Princeton University Press, New Jersey.

——— (2008): "Financial Markets and the Real Economy," in *Handbook of the Equity Risk Premium*. Elsevier.

Cochrane, J. H., and M. Piazzesi (2005): "Bond Risk Premia," *American Economic Review*, 95(1), 138–160.

Cooper, I., and R. Priestley (2011): "Real Investment and Risk Dynamics," *Journal of Financial Economics*, 101, 182–205.

Cooper, M. J., R. C. J. Gutierrez, and A. Hameed (2004): "Market States and Momentum," *Journal of Finance*, 59, 1345–1365.

Dales, A., and R. Meese (2001): "Strategic Currency Hedging," *Journal of Asset Management*, 2, 9–21.

Davis, M. A., and J. Heathcote (2005): "Housing and the Business Cycle," *International Economic Review*, 46, 751–784.

de Roon, F. A., and T. Nijman (2001): "Testing Mean-Variance Spanning: A Survey," *Journal of Empirical Finance*, 8, 111–155.

de Roon, F. A., T. Nijman, and B. J. M. Werker (2003): "Currency Hedging for International Stock Portfolios: The Usefulness of Mean-Variance Analysis," *Journal of Banking and Finance*, 27, 327–349.

de Santis, G., and B. Gerard (1997): "International Asset Pricing and Portfolio Diversification with Time-Varying Risk," *Journal of Finance*, 52, 1881–1912.

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, 1915–1953.

den Haan, W. J., and A. T. Levin (2000): "Robust Covariance Matrix Estimation with Data-Dependent VAR Prewhitening Order," NBER Technical Working Paper 255.

Driessen, J., and L. Laeven (2007): "International Portfolio Diversification Benefits: Cross-Country Evidence from a Local Perspective," *Journal of Banking and Finance*, 31, 1693–1712.

Eichholtz, P. M. A. (1996): "Does International Diversification Work Better for Real Estate than for Stocks and Bonds?," *Financial Analysts Journal*, 52, 56–62.

Engsted, T., and S. V. Møller (2011): "Cross-sectional Consumption-based Asset Pricing: the Importance of Consumption Timing and the Inclusion of Severe Crises," CREATES Research Paper.

Errunza, V., K. Hogan, and M.-W. Hung (1999): "Can the Gains from International Diversification be Achieved without Trading Abroad?," *Journal of Finance*, 54, 2075–2107.

Eun, C. S., W. Huang, and S. Lai (2008): "International Diversification with Large- and Small-Cap Stocks," *Journal of Financial and Quantitative Analysis*, 43, 489–523.

Eun, C. S., S. Lai, F. A. de Roon, and Z. Zhang (2010): "International Diversification with Factor Funds," *Management Science*, 56, 1500–1518.

Eun, C. S., and J. Lee (2010a): "Mean-Variance Convergence around the World," *Journal of Banking and Finance*, 34, 856–870.

——— (2010b): "Mean-Variance Convergence Around the World," *Journal of Banking & Finance*, 34, 856–870.

Eun, C. S., and B. G. Resnick (1988): "Exchange Rate Uncertainty, Forward Contracts, and International Portfolio Selection," *Journal of Finance*, 43, 197–216.

Fama, E. F. (1984): "Forward and Spot Exchange Rates," *Journal of Monetary Economics*, 14, 319–338.

——— (1996): "Multifactor Portfolio Efficiency and Multifactor Asset Pricing," *Journal of Financial and Quantitative Analysis*, 31, 441–465.

Fama, E. F., and K. R. French (1992): "The Cross-Section of Expected Stock Returns," *Journal of Finance*, 47(2), 427–465.

——— (1993): "Common Risk Factors in the Returns on Stocks and Bonds," *Journal of Financial Economics*, 33, 3–56.

——— (1996): "Multifactor Explanations of Asset Pricing Anomalies," *Journal of Finance*, 51, 55–84.

——— (1998): "Value versus Growth: The International Evidence," *Journal of Finance*, 53, 1975–1999.

Fama, E. F., and J. D. MacBeth (1973): "Risk, Return, and Equilibrium: Emprical Tests," *Journal of Political Economy*, 81, 607–636.

Farhi, E., S. P. Fraiberger, X. Gabaix, R. Rancière, and A. Verdelhan (2009): "Crash Risk in Currency Markets," CEPR Discussion Paper.

Ferguson, M. F., and R. L. Shockley (2003): “Equilibirum ”Anomalies”,” *Journal of Finance*, 58, 2549–2580.

Ferson, W., A. F. Siegel, and P. Xu (2006): “Mimicking Portfolios with Conditioning Information,” *Journal of Financial and Quantitative Analysis*, 41, 607–635.

Ferson, W. E., and C. R. Harvey (1992): “Seasonality and Consumption-based Asset Pricing,” *Journal of Finance*, 47, 511–552.

Ferson, W. E., and A. F. Siegel (2010): “Testing Portfolio Efficiency with Conditioning Information,” *Review of Financial Studies*, 22, 2735–2758.

Fisher, J. D. (2007): “Why Does Household Investment Lead Business Investment over the Business Cycle,” *Journal of Political Economy*, 115, 188–214.

Froot, K. A., and R. Thaler (1990): “Anomalies Foreign Exchange,” *Journal of Economic Perspectives*, 4(3), 179–192.

Galati, G., A. Heath, and P. McGuire (2007): “Evidence of Carry Trade Activity,” *BIS Quarterly Review*, (Sept.), 27–41.

Garman, M. B., and S. W. Kohlhagen (1983): “Foreign Currency Option Values,” *Journal of International Money and Finance*, 2, 231–237.

Geltner, D. (1989): “Estimating Real Estate’s Systematic Risk from Aggregate Level Appraisal-based Returns,” *AREUEA Journal*, 17, 463–481.

——— (1993): “Estimating Market Values from Appraised Values without Assuming an Efficient Market,” *Journal of Real Estate Research*, 8, 325–345.

Gibbons, M. R., S. A. Ross, and J. Shanken (1989): “A Test of the Efficiency of a Given Portfolio,” *Econometrica*, 57(5), 1121–1152.

Gilmore, S., and F. Hayashi (2011): "Emerging Market Currency Excess Returns," *American Economic Journal: Macroeconomics*, 3, 85–111.

Glabadanidis, P. (2009): "Measuring the Economic Significance of Mean-Variance Spanning," *Quarterly Review of Economics and Finance*, 49, 596–616.

Glen, J., and P. Jorion (1993): "Currency Hedging for International Portfolios," *Journal of Finance*, 48, 1865–1886.

Goetzmann, W. N., L. Li, and K. G. Rouwenhorst (2005): "Long-Term Global Market Correlations," *Journal of Business*, 78, 1–38.

Gomme, P., F. E. Kydland, and P. Rupert (2001): "Home Production Meets Time to Build," *Journal of Political Economy*, 109, 115–131.

Gorton, G. B., F. Hayashi, and K. G. Rouwenhorst (2012): "The Fundamentals of Commodity Futures Returns," *Review of Finance*, 17, 35–105.

Greenwood, J., and Z. Hercowitz (1991): "The Allocation of Capital and Time over the Business Cycle," *Journal of Political Economy*, 99, 188–214.

Gregory, A. W., and T. Wirjanto (1993): "Discussion: The Effect of Sampling Error on the Time Series Behavior of Consumption Data," *Journal of Econometrics*, 55, 267–273.

Griffin, J., and J. Martin (2003): "Momentum Investing and Business Cycle Risk: Evidence from Pole to Pole," *Journal of Finance*, 58, 2515–2547.

Grossman, S. J., A. Melino, and A. Shiller (1987): "Estimating the Continuous-Time Consumption-based Asset Pricing Model," *Journal of Business and Economic Statistics*, 5, 315–327.

Grubel, H. G. (1968): "Internationally Diversified Portfolios: Welfare Gains and Capital Flows," *American Economic Review*, 58, 1299–1314.

GYNTELBERG, J., AND E. M. REMOLONA (2007): "Risk in Carry Trades: A Look at Target Currencies in Asia and the Pacific," *BIS Quarterly Review*, (Dec.), 73–82.

HALL, R. E. (1978): "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence," *Journal of Political Economy*, 86, 971–987.

——— (1988): "Intertemporal Substitution in Consumption," *Journal of Political Economy*, 96, 339–357.

HANSEN, L. P., AND R. JAGANNATHAN (1991): "Implications of Security Market Data for Models of Dynamic Economies," *Journal of Political Economy*, 99, 225–262.

——— (1997): "Assessing Specification Errors in Stochastic Discount Factor Models," *Journal of Finance*, 52, 557–590.

HANSEN, L. P., AND K. H. SINGLETON (1982): "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models," *Econometrica*, 50, 1269–1286.

HARVEY, C. R. (1995): "Predictable Risk and Returns in Emerging Markets," *Review of Financial Studies*, 8, 773–816.

HAYASHI, F., AND C. SIMS (1983): "Nearly Efficient Estimation of Time Series Models with Predetermined, but not Exogenous Instruments," *Econometrica*, 51, 783–798.

HESTON, A., R. SUMMERS, AND B. ATEN (2009): "Penn World Table, Income and Prices at the University of Pennsylvania," Center for International Comparisons of Production, Version 6.3, August.

HUBERMAN, G., AND S. KANDEL (1987): "Mean-Variance Spanning," *Journal of Finance*, 42, 873–888.

JAGANNATHAN, R., S. MARAKANI, H. TAKEHARA, AND Y. WANG (2012): "Calender Cycles, Infrequent Decisions, and the Cross-Section of Stock Returns," *Management Science*, 58, 507–522.

Jagannathan, R., and Y. Wang (2007): "Lazy Investors, Discretionary Consumption, and the Cross-Section of Stock Returns," *Journal of Finance*, 62, 1623–1661.

Jagannathan, R., and Z. Wang (1996): "The Conditional CAPM and the Cross-Section of Expected Returns," *Journal of Finance*, 51, 3.53.

Jegadeesh, N., and S. Titman (1993): "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency," *Journal of Finance*, 48, 65–91.

Jobson, J. D., and B. M. Korkie (1989): "A Performance Interpretation of Multivariate Tests of Asset Set Intersection, Spanning, and Mean-Variance Efficiency," *Journal of Financial and Quantitative Analysis*, 24, 185–204.

Jorda, O., and A. M. Taylor (2012): "The Carry Trade and Fundamentals: Nothing to fear but FEER itself," *Journal of International Economics*, 88, 74–90.

Jorion, P. (1994): "Mean/Variance Analysis of Currency Overlays," *Financial Analysts Journal*, 50, 48–56.

Jurek, J. W. (2009): "Crash-neutral Currency Carry Trades," Working Paper.

Kalman, R. E. (1960): "A New Approach to Linear Filtering and Prediction Problems," *Transactions of the ASME–Journal of Basic Engineering*, 82, 35–45.

Kan, R., and G. Zhou (2012a): "Tests of Mean-Variance Spanning," *Annals of Economics and Finance*, 13, 145–193.

——— (2012b): "Tests of Mean-Variance Spanning," *Annals of Economics and Finance*, 13, 145–193.

Kandel, S., and R. F. Stambaugh (1995): "Portfolio Inefficiency and the Cross-Section of Expected Returns," *Journal of Finance*, 50, 157–184.

Koijen, R. S., H. Lustig, and S. V. Nieuwerburgh (2012): "The Cross-Section and Time-Series of Stock and Bond Returns," June 2012.

Lamont, O. A. (2001): "Economic Tracking Portfolios," *Journal of Econometrics*, 105, 161–184.

Leamer, E. E. (2007): "Housing is the Business Cycle," Housing, Housing Finance and Monetary Policy, A Symposium Sponsored by the Federal Reserve of Kansas City.

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

Lettau, M., and S. Ludvigson (2001): "Resurrecting the (C)CAPM: A Cross-Sectional Test when Risk Premia Are Time-Varying," *Journal of Political Economy*, 109, 1238–1287.

Lettau, M., and S. C. Ludvigson (2009): "Euler Equation Errors," *Review of Economic Dynamics*, 12, 255–283.

Levy, H. (2006): *Stochastic Dominance*. Springer Science Business Media.

Lewellen, J., S. Nagel, and J. Shanken (2010): "A Skeptical Appraisal of Asset Pricing Tests," *Journal of Financial*, 96, 175–194.

Li, Q., M. Vassalou, and Y. Xing (2006): "Sector Investment Growth Rates and the Cross Section of Equity Returns," *Journal of Business*, 79, 1637–1665.

Liu, C. H., and J. Mei (1998): "The Predictability of International Real Estate Markets, Exchange Rate Risks and Diversification Consequences," *Real Estate Economics*, 26, 3–39.

Liu, L. X., and L. Zhang (2008): "Momentum Profits, Factor Pricing, and Macroeconomic Risk," *Review of Financial Studies*, 21, 2417–2448.

Ljungqvist, L., and T. J. Sargent (2004): "Recursive Macroeconomic Theory," MIT Press, Cambridge.

Longin, F., and B. Solnik (1995): "Is the Correlation in International Equity Returns Constant: 1960-1990?," *Journal of International Money and Finance*, 14, 3–26.

——— (2001): "Extreme Correlation on International Equity Markets," *Journal of Finance*, 56, 649–676.

Lucas, R. E. (1978): "Asset Prices in an Exchange Economy," *Econometrica*, 46, 1429–1445.

Ludvigson, S. C. (2012): "Advances in Consumption-based Asset Pricing: Empirical Tests," *Handbook of the Economics of Finance, edited by George Constantinides, Milton Harris and Rene Stulz*, Volume 2, forthcoming.

Lustig, H., N. Roussanov, and A. Verdelhan (2011): "Common Risk Factors in Currency Markets," *Review of Financial Studies*, 24, 3731–3777.

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

Lyons, R. K. (2001): *The Microstructure Approach to Exchange Rates.* Cambridge: MIT Press.

Maio, P., and P. Santa-Clara (2012): "Multifactor Models and their Consistency with the ICAPM," *Journal of Financial Economics*, forthcoming.

Mehra, R., and E. C. Prescott (1985): "The Equity Premium: A Puzzle," *Journal of Monetary Economics*, 15, 145–161.

Menkhoff, L., L. Sarno, M. Schmeling, and A. Schrimpf (2012a): "Carry Trades and Global Foreign Exchange Volatility," *Journal of Finance*, 67, 681–718.

——— (2012b): "Currency Momentum Strategies," *Jounal of Financial Economics*, 106, 660–684.

Menkhoff, L., and M. P. Taylor (2007): "The Obstinate Passion of Foreign Exchange Professionals: Technical Analysis," *Journal of Economic Literature*, 45, 936–972.

Newey, W. K., and K. D. West (1987): "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica,*, 55, 703–708.

Oikarinen, E., M. Hoesli, and C. Serrano (2011): "The Long-run Dynamics Between Direct and Securitized Real Estate," *Journal of Real Estate Research*, 33, 73–103.

Okunev, J., and D. White (2003): "Do Momentum-Based Strategies Still Work in Foreign Currency Markets?," *Journal of Financial and Quantitative Analysis*, 38, 425–447.

Olma, A. R., and L. B. Siegel (2004): "A New Framework for International Investing," *Journal of Portfolio Management*, 30, 55–69.

Palazzo, G., and S. Nobili (2010): "Explaining and Forecasting Bond Risk Premiums," *Financial Analysts Journal*, 66, 67–82.

Parker, J. A., and C. Julliard (2005): "Consumption Risk and the Cross Section of Expected Returns," *Journal of Political Economy*, 113, 185–222.

Perold, A. F., and E. C. Schulman (1988): "The Free Lunch in Currency Hedging: Implications for Investment Policy and Performance Standards," *Financial Analysts Journal*, 44, 45–50.

Petkova, R. (2006): "Do the Fama-French Factors Proxy for Innovations in Predictive Variables?," *Journal of Finance*, 61, 581–612.

Pojarliev, M., and R. M. Levich (2008): "Do Professional Currency Managers Beat the Benchmark?," *Financial Analysts Journal*, September-October 64, 18–32.

——— (2012): *Is There Skill or Alpha in Currency Investing? In: The Handbook of Exchange Rates, Jessica James, Ian March and Lucio Sarno (eds.)*. Wiley.

Post, T., and P. Versijp (2007): "Multivariate Tests for Stochastic Dominance Efficiency of a Given Portfolio," *Journal of Financial and Quantitative Analysis*, 42, 489–515.

Quan, D. C., and J. M. Quigley (1989): “Inferring an Investment Return Series for Real Estate from Observations on Sales,” *AREUEA Journal*, 17, 218–234.

Rouwenhorst, K. G. (1998): “International Momentum Strategies,” *Journal of Finance*, 53, 267–284.

Rubens, J. H., D. A. Louton, and E. J. Yobaccio (1998): “Measuring the Significance of Diversification Gains,” *Journal of Real Estate Research*, 16, 73–86.

Sa-Aadu, J., J. Shilling, and A. Tiwari (2010): “On the Portfolio Properties of Real Estate in Good Times and Bad Times,” *Real Estate Economics*, 38, 529–565.

Sadka, R., and R. A. Korajczyk (2004): “Are momentum profits robust to trading costs?,” *Journal of Finance*, 59, 1039–1082.

Sarno, L., P. D. Corte, and I. Tsiakas (2011): “Spot and Forward Volatility in Foreign Exchange,” *Journal of Financial Economics*, 100, 496–513.

Savov, A. (2011): “Asset Pricing with Garbage,” *Journal of Finance*, 66, 177–201.

Serrano, C., and M. Hoesli (2009): “Global Securitized Real Estate Benchmarks and Performance,” *Journal of Real Estate Portfolio Management*, 15, 1–19.

Shanken, J. (1992): “On the Estimation of Beta-Pricing Models,” *Review of Financial Studies*, 5, 1–33.

Solnik, B. H. (1974): “Why Not Diversify Internationally Rather than Domestically?,” *Financial Analysts Journal*, 30, 48–54.

Summers, L. H. (1985): “On Economics and Finance,” *Journal of Finance*, 40, 633–635.

Taio, G. C. (1972): “Asymptotic Behaviour of Temporal Aggregates of Time Series,” *Biometrika*, 59, 525–531.

Tsay, R. S. (2005): “Analysis of Financial Time Series,” Wiley-Interscience, New Jersey.

VASSALOU, M. (2003): "News Related to Future GDP Growth as a Risk Factor in Equity Returns," *Journal of Financial Economics*, 68, 47–73.

VERDELHAN, A. (2010): "A Habit-Based Explanation of the Exchange Rate Risk Premium," *Journal of Finance*, 65, 123–146.

——— (2011): "The Share of Systematic Variation in Bilateral Exchange Rates," Working Paper, November 2011.

WEIL, P. (1989): "The Equity Premium Puzzle and the Risk-Free Rate Puzzle," *Journal of Monetary Economics*, 24, 401–422.

WILCOX, D. W. (1992): "The Construction of U.S. Consumption Data: Some Facts and Their Implications for Empirical Work," *Amercian Economic Review*, 82, 922–941.

WORKING, H. (1960): "Note on the Correlation of First Differences of Averages in a Random Chain," *Econometrica*, 28, 916–918.

WORZALA, E., AND C. F. SIRMANS (2003): "Investing in International Real Estate Stocks: A Review of the Literature," *Urban Studies*, 40, 1115–1149.

YOGO, M. (2006): "A Consumption-based Explanation of Expected Stock Returns," *Journal of Finance*, 61, 539–580.

YUNUS, N., J. A. HANSZ, AND P. J. KENNEDY (2012): "Dynamic Interactions Between Private and Public Real Estate Markets: Some International Evidence," *Journal of Real Estate Finance and Economics*, forthcoming.

Appendix

A Asset Pricing without Garbage

A.1 Simulation: Details on the Asset Pricing Economy

Model. This section provides details on how I generate the data of the model economy in which asset pricing does work. The Euler equation satisfies the pricing restriction:

$$E_t\left[exp\left(ln(\beta) - \gamma \triangle c_{t+1} + r_{m,t+1}\right)\right] = 1,$$

where $r_{m,t+1} = ln\left(R_{M,t+1}\right)$ is the log stock market return and $\triangle c_{t+1} = ln(C_{t+1}/C_t)$ is log consumption growth. Both are driven by the following dynamics:

$$\triangle c_{t+1} = \mu + \sigma \eta_{t+1},$$

$$\triangle d_{t+1} = \mu_d + \varphi_d \sigma u_{t+1} + \pi_d \sigma \eta_{t+1},$$

$$u_{t+1}, \eta_{t+1} \sim N\left(0, 1\right),$$

where $\triangle d_{t+1}$ is log dividend growth. This model is the canonical CCAPM case of Bansal and Yaron (2004), and can be solved for the process of the stock market return and the risk-free rate as described in their paper. In the particular, the log stock market return at time $t+1$ is given by:

$$r_{m,t+1} = -ln(\beta) + \gamma\mu - \frac{1}{2}\left[\left(\pi_d - \gamma\right)^2\sigma^2 + \varphi_d\sigma^2\right] + \varphi_d\sigma u_{t+1} + \pi_d\sigma\eta_{t+1},$$

and the implied log risk-free interest rate is:

$$r_f = -ln(\beta) + \gamma\mu - \frac{1}{2}\gamma^2\sigma^2.$$

For a given set of parameter values $(\beta, \gamma, \mu, \sigma, \mu_d, \varphi_d, \pi_d)$, I can simulate consumption growth and the model implied equity premium, $r_{m,t+1} - r_f$.

Parameters. Table A.1 reports the parameter values which I use to simulate monthly data. These parameters are calibrated such that simulated ”true” consumption growth has similar characteristics as garbage. Benchmarking to garbage is useful, since we know from Savov (2011) that the stochastic properties of garbage can match the equity premium in the data quite well.

The empirical correlation between the stock market and garbage is about 0.60 and the annual standard deviation of garbage is 3%. I set the dividend consumption exposure parameter to 3.8 and consumption volatility to the according annual value of 3% to generate similar properties for true consumption growth in the model economy. The coefficient of relative risk aversion of 15 together with the other parameters implies an annual expected equity premium of 5.1%.[74]

[74] The expected equity risk premium in this model economy is $E\left(r_{m,t+1} - r_f\right) + \frac{1}{2}Var\left(r_{m,t+1}\right) = \gamma\pi_d\sigma^2$.

Finally, I set the dividend leverage parameter to 5.5 to generate annual stock market volatility of 20%. The time discount factor, the mean of consumption growth, and the mean of dividend growth do not effect the equity premium and are set to common values in the related literature.

Table A.1: Parameter Values for the Simulated Asset Pricing Economy

	Variable	Value
Risk Aversion	γ	15
Time Discount Factor	β	$0.988^{1/12}$
Mean Consumption Growth	μ	0.01/12
Consumption Volatility	σ	$0.03/(12^{1/2})$
Mean Dividend Growth	μ_d	0.01/12
Dividend Consumption Exposure	π_d	3.8
Dividend Leverage	φ_d	5.5

Adding Time Aggregation and Filtering. I convert simulated monthly log consumption growth rates and log returns into simple growth rates and simple returns, and also produce a series of monthly consumption levels. The table in the main paper display statistics for the annual simple growth rates of consumption, to be comparable to the empirical part of the paper. Time aggregation and filtering are then added to the data as follows: The subscript consists of a month index m running from one to twelve and a year index t running from zero to T. The year $t+1$ consumption growth rate is based on "December" levels:

$$\triangle C_{t+1} = C_{12;t+1}/C_{12;t} - 1. \tag{37}$$

The year $t+1$ consumption growth rate with time aggregation ("TA") is calculated as:

$$\triangle c_{t+1}^{TA} = log(\sum_{m=1}^{12} C_{m;t+1} / \sum_{m=1}^{12} C_{m;t}). \tag{38}$$

To simulate a filter, I assume that the variance of (individual consumption component) measurement error is as large as the variance of true consumption growth, which gives $\nu = 1/2$. For aggregate consumption, measurement error diversifies and is infinitesimal small and only the filter carries over. Thus I directly apply the simulated filter on aggregate log consumption:

$$c_{t+1}^{TAF} = 1/2 \times c_{t+1}^{TA} + (1 - 1/2) \times c_t^{TAF}, \tag{39}$$

$$\triangle c_{t+1}^{TAF} = c_{t+1}^{TAF} - c_t^{TAF}$$

The resulting year $t+1$ consumption is subject to time aggregation and filtering ("TAF") and corresponds to consumption as reported in NIPA. The equity premium is based on "December" levels, and is used to calculate the correlations and covariances of all investigated consumption measures.

Unfiltering. I compare three different "unfilter" rules for "time aggregated" and "time aggregated and filtered" consumption. First, I apply the new method to the simulated data:

$$\triangle c_{t+1}^{\phi} = \left[\triangle c_{t+1}^{n} - (1-\phi)\,\triangle c_{t}^{n}\right]/\phi, \;\; n = TA,\, TAF, \tag{40}$$

$$\triangle C_{t+1}^{\phi} = exp(1 + \triangle c_{t+1}^{\phi}) - 1.$$

Since the data-generating process and the used filter are known within the simulation experiment, it is possible to anticipate an unfilter parameter which should yield a good result matching "true" consumption growth. Using the approximation $\phi = .80 \times \nu$ gives a parameter of $\phi = .80$ if the data is not filtered ($\nu = 1$), and of $\phi = .40$ if the data is filtered ($\nu = 1/2$). As a sensitivity check, results are presented for ϕ equal to .80, .53, and .40, for both, "time aggregated" and "time aggregated and filtered" consumption.

Second, I consider three-year consumption (P-J) following Parker and Julliard (2005):

$$\triangle C_{t+1}^{P-J} = C_{t+3}^{n}/C_{t}^{n} - 1, \;\; n = TA,\, TAF. \tag{41}$$

Third, I calculate fourth-quarter to fourth-quarter consumption (Q4Q4) as in Jagannathan and Wang (2007):

$$\triangle C_{t+1}^{Q4Q4} = C_{Q4;t+1}^{n}/C_{Q4;t}^{n} - 1, \;\; n = TA,\, TAF, \tag{42}$$

where $C_{Q4;t}^{n} = \sum_{m=10}^{12} C_{m;t}^{n}$. The monthly simulated consumption process is used to calculate the Q4Q4 measure if there is only time aggregation in consumption data ($n = TA$). For "time aggregated and filtered" consumption ($n = TAF$), I apply the proportional denton method with the annual filtered level of consumption to benchmark monthly consumption, which results in monthly and quarterly filtered consumption. The proportional denton method is actually used by the BEA to benchmark monthly consumption data on annual consumption data (see BEA (2009) for details).

Simulation Procedure. I simulate the model economy 10,000 times with 792 observations. The first "year" (i.e., the first twelve) observations are needed to calculate annually time aggregated consumption levels for the first year. The next five years of observations are taken as "burn-in" period to calculate filtered consumption. Thus, in each run, 60 years of annual observations on growth rates and returns are included for further analysis.

A.2 Empirical Data

Consumption. Annual NIPA consumption data is collected from the NIPA tables 2.3.4 / 2.3.5 for "nondurable goods" (Line 8) and "services" (Line 13). The corresponding price indices and the population reported in NIPA table 7.1 (Line 18) are used to compute real per capita growth rates.

NIPA consumption growth data starts in 1930. However, stock market data available on the web site of Kenneth R. French begins in 1927. To not just cut off the stock market crash of 1929 from the sample, I use the consumption measure (including nondurables and services) provided by Robert Shiller on his web site for the period from 1926 to 1929 (which is based on the

NBER Kendricks consumption expenditure series), and splice the series with NIPA consumption growth. The NBER Kendricks consumption expenditure series closely follows the measurement methodology used for NIPA consumption. However, there is no consumption measure available for 1926 to 1929 which measures nondurables and services separately. Thus, I assume that the real per capita growth rate for these four years is identical for nondurables and services and splice the series with NIPA nondurables and services and NIPA nondurables. Real per capita consumption growth for 1926 is needed to construct the unfiltered NIPA consumption growth of 1927.

Financial Returns. Annual stock market returns are collected from the web site of Kenneth R. French (the market return, and 25 Fama-French portfolios sorted by size and book-to-market). I use the one-year risk-free interest rate provided on the web site of Robert Shiller to calculate the equity premium and excess returns on the 25 Fama-French portfolio returns. Stock market returns are deflated using the inflation rate implied by the consumption measure (NIPA nondurables and services).

Timing Conventions. If consumption is measured as the flow of consumption during the interval of a specific period, a timing convention may be used to attribute consumption to one point of time, e.g. to the end or the beginning of the considered period, as suggested by Campbell (2003). Campbell finds that the stock market correlation using the "beginning-of-period" timing convention (i.e. matching stock market returns with one period lagged NIPA consumption growth) is larger than using the "end-of-period" timing convention (i.e. matching contemporaneous stock market returns with NIPA consumption growth).

B GDP Mimicking Portfolios and the Cross-Section of Stock Returns

B.1 Estimation without Intercept

The inclusion or exclusion of a constant (common pricing error) can have a substantial impact on cross-sectional regression results (see e.g. Burnside (2011)). However, Table B.1 shows that we can draw the same conclusions from estimation without an intercept.

B.2 Growth Rates of Components

Table B.2 documents that our findings are robust to the use of growth rates of aggregate GDP and GDP components instead of the mimicking portfolios. We also present estimates without intercept in Table B.3.

B.3 Additional Results

In the main text, we base our analyzes on the empirical fact that some GDP components lead aggregate GDP and some GDP components lag aggregate GDP. This stylized pattern of the data can be also illuminated using the event-based approach proposed by Koijen, Lustig, and Nieuwerburgh (2012).

First, we define an event as the lowest 30% realizations of a specific time series, e.g. aggregate GDP, and label this event '0'. Second, we pick dates prior ('-1', '-2', '-3') and dates following ('1', '2, '3') the event. Third, we plot the average realization of a specific variable (e.g., growth of residential investment) during this event time line.

Lead and Lag in GDP Components. Figure B.1 provides annual real per capita aggregate GDP and the five GDP components in aggregate GDP-event time. For comparisons, Figure B.2 shows aggregate GDP and the five GDP components during NBER recessions. We find in both figures that the lead-lag behavior of GDP components shows up.

Forecasting Regressions with Non-standardized GDP. Table B.4 provides forecasting regressions of aggregate GDP on one period lagged components of GDP. In difference to the main text, we do not standardize the variables.

Table B.1: Monthly GDP Mimicking Portfolio Factors - Estimation without Intercept

The table reports GMM estimates of factor prices ($\boldsymbol{\lambda}$) of the monthly GDP mimicking portfolio factors as in the main paper but with a zero constant ($\gamma = 0$):

$$E\left(\mathbf{R}_t\right) = 0 + \boldsymbol{\beta}\boldsymbol{\lambda}.$$

T-statistics are reported in parentheses and are based on the Newey and West (1987) HAC covariance matrix with automatic lag length selection according to Andrews (1991). All data are monthly and the sample period is from January 1951 to December 2010.

	γ	Factor Price ($\boldsymbol{\lambda}$)		R^{2OLS}	γ	Factor Price ($\boldsymbol{\lambda}$)		R^{2OLS}	γ	Factor Price ($\boldsymbol{\lambda}$)			R^{2OLS}
	FF25				Momentum 10				FF25 + M10				
(1)	γ	MKT	RES		γ	MKT	RES		γ	MKT	RES		
	-	0.59	1.31	0.74	-	0.57	1.92	0.86	-	0.57	1.50		0.78
		(3.12)	(5.20)			(3.17)	(4.18)			(3.07)	(7.82)		
(2)	γ	MKT	DUR		γ	MKT	DUR		γ	MKT	DUR		
	-	0.73	2.37	0.72	-	0.62	1.63	0.88	-	0.70	1.93		0.77
		(3.75)	(4.19)			(3.49)	(3.97)			(3.75)	(6.20)		
(3)	γ	MKT	NDU		γ	MKT	NDU		γ	MKT	NDU		
	-	0.73	1.02	0.26	-	0.61	0.87	0.88	-	0.70	0.95		0.60
		(3.78)	(3.26)			(3.46)	(4.61)			(3.73)	(6.68)		
(4)	γ	MKT	EQS		γ	MKT	EQS		γ	MKT	EQS		
	-	0.79	0.90	-0.46	-	0.64	0.96	0.87	-	0.75	0.95		0.26
		(4.52)	(1.04)			(3.63)	(3.37)			(4.11)	(3.43)		
(5)	γ	MKT	BST		γ	MKT	BST		γ	MKT	BST		
	-	0.70	-0.38	-0.56	-	0.63	0.95	0.88	-	0.74	0.66		0.09
		(3.92)	(-0.62)			(3.61)	(3.15)			(4.09)	(2.42)		
(6)									γ	MKT	RES	BST	
									-	0.61	1.45	0.54	0.82
										(3.43)	(7.32)	(1.93)	

Table B.2: Annual Macroeconomic Variables

The table reports GMM estimates of a constant (γ) and factor prices ($\boldsymbol{\lambda}$) as in the main paper but we use annual real per capita growth rates of GDP components instead of the mimicking portfolios:

$$E\left(\mathbf{R}_t\right) = \gamma + \boldsymbol{\beta}\boldsymbol{\lambda}.$$

T-statistics are reported in parentheses and are based on the Newey and West (1987) HAC covariance matrix with automatic lag length selection according to Andrews (1991). All data are annual and the sample period is from 1951 to 2010.

	γ	Factor Price ($\boldsymbol{\lambda}$)		R^{2OLS}	γ	Factor Price ($\boldsymbol{\lambda}$)		R^{2OLS}	γ	Factor Price ($\boldsymbol{\lambda}$)			R^{2OLS}
	FF25				Momentum 10				FF25 + M10				
(1)	γ	MKT	RES		γ	MKT	RES		γ	MKT	RES		
	5.75	3.12	17.48	0.74	11.34	-2.78	28.16	0.39	8.19	0.50	19.18		0.57
	(0.87)	(0.47)	(2.49)		(1.06)	(-0.25)	(1.57)		(1.24)	(0.08)	(2.46)		
(2)	γ	MKT	DUR		γ	MKT	DUR		γ	MKT	DUR		
	4.92	5.34	10.83	0.48	8.27	0.82	17.11	0.76	5.66	4.28	13.52		0.53
	(0.67)	(0.71)	(2.13)		(1.12)	(0.11)	(1.52)		(0.80)	(0.58)	(3.03)		
(3)	γ	MKT	NDU		γ	MKT	NDU		γ	MKT	NDU		
	5.87	4.24	1.66	0.25	3.99	5.35	2.42	0.89	3.88	5.96	2.27		0.58
	(1.21)	(0.87)	(1.68)		(0.71)	(0.93)	(1.63)		(0.69)	(1.03)	(1.82)		
(4)	γ	MKT	EQS		γ	MKT	EQS		γ	MKT	EQS		
	17.47	-9.05	-9.44	0.21	6.96	2.41	11.33	0.96	10.04	0.46	5.69		0.15
	(2.22)	(-1.24)	(-1.33)		(1.68)	(0.54)	(2.00)		(2.38)	(0.10)	(1.84)		
(5)	γ	MKT	BST		γ	MKT	BST		γ	MKT	BST		
	20.63	-12.19	-10.89	0.54	1.21	8.05	10.32	0.90	11.50	-1.62	2.41		0.06
	(2.08)	(-1.26)	(-1.95)		(0.26)	(1.42)	(1.86)		(2.69)	(-0.34)	(0.97)		
(6)									γ	MKT	RES	BST	
									3.63	5.59	20.72	3.40	0.64
									(0.54)	(0.76)	(2.13)	(0.71)	

Table B.3: Annual Macroeconomic Variables - Estimation without Intercept

The table reports GMM estimates of factor prices ($\boldsymbol{\lambda}$) as in the main paper but we use annual real per capita growth rates of GDP components and impose a zero constant ($\gamma = 0$):

$$E\left(\mathbf{R}_t\right) = 0 + \boldsymbol{\beta}\boldsymbol{\lambda}.$$

T-statistics are reported in parentheses and are based on the Newey and West (1987) HAC covariance matrix with automatic lag length selection according to Andrews (1991). All data are annual and the sample period is from 1951 to 2010.

	γ	Factor Price ($\boldsymbol{\lambda}$)		R^{2OLS}	γ	Factor Price ($\boldsymbol{\lambda}$)		R^{2OLS}	γ	Factor Price ($\boldsymbol{\lambda}$)			R^{2OLS}
	FF25				Momentum 10				FF25 + M10				
(1)	γ	MKT	RES		γ	MKT	RES		γ	MKT	RES		
	-	8.29	20.01	0.66	-	8.39	43.61	0.19	-	7.97	22.94		0.45
		(4.11)	(2.66)			(3.67)	(1.16)			(3.99)	(2.70)		
(2)	γ	MKT	DUR		γ	MKT	DUR		γ	MKT	DUR		
	-	9.95	12.44	0.43	-	8.93	20.52	0.65	-	9.64	15.44		0.48
		(3.21)	(1.89)			(4.44)	(1.64)			(3.70)	(2.89)		
(3)	γ	MKT	NDU		γ	MKT	NDU		γ	MKT	NDU		
	-	9.73	2.12	0.19	-	9.33	2.67	0.87	-	9.64	2.51		0.56
		(4.03)	(1.94)			(4.87)	(1.65)			(4.70)	(2.08)		
(4)	γ	MKT	EQS		γ	MKT	EQS		γ	MKT	EQS		
	-	8.93	-2.12	-0.37	-	9.24	12.37	0.88	-	10.39	7.96		-0.01
		(3.93)	(-0.52)			(4.93)	(2.07)			(5.08)	(2.43)		
(5)	γ	MKT	BST		γ	MKT	BST		γ	MKT	BST		
	-	8.27	-5.14	-0.28	-	9.27	10.57	0.90	-	9.76	5.09		-0.12
		(3.99)	(-1.27)			(3.94)	(1.76)			(4.17)	(1.79)		
(6)									γ	MKT	RES	BST	
									-	9.15	22.25	4.20	0.63
										(4.55)	(2.50)	(0.95)	

Figure B.1: Low Aggregate GDP Events

The figure shows annual real per capita growth of aggregate GDP and five GDP components in aggregate GDP event time. The event is defined as years with the 30% lowest realizations of aggregate GDP growth and is labeled '0'. The years '-1', '-2, and '-3' refer to one, two, and three years before the event takes place. The years '1', '2, and '3' refer to one, two, and three years after the event takes place. The panels plot aggregate GDP (GDP), residential investment (RES), durables (DUR), nondurables (NDU), equipment and software (EQS) and business structures (BST), and the sample is from 1951 to 2010.

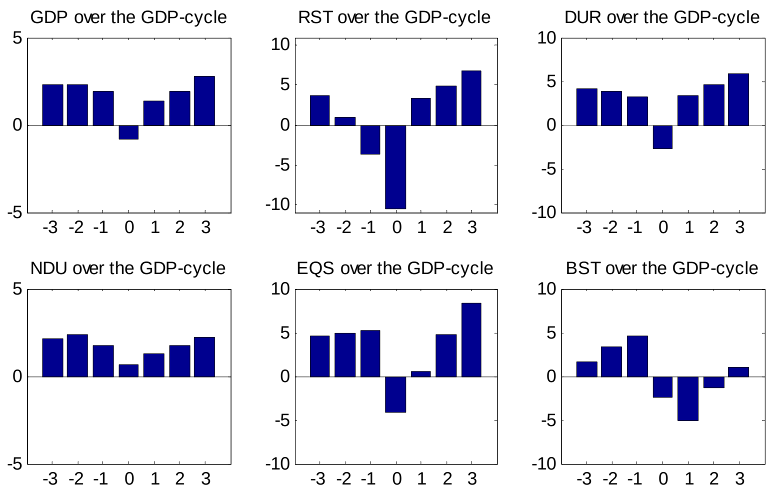


Figure B.2: Recession Events

The figure shows annual real per capita growth of aggregate GDP and five GDP components during recessions. The beginning of a recession year is defined as a year with more than two recession quarters according to NBER and is labeled '0'. The sample is from 1951 to 2010.

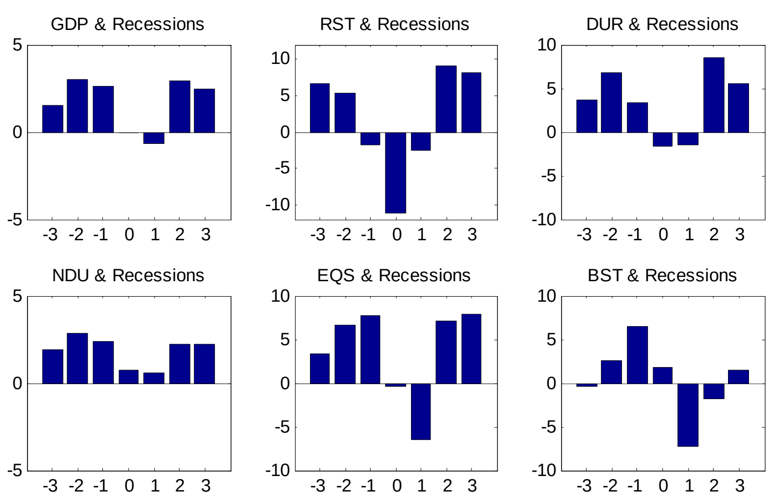


Table B.4: Forecasting Regressions with Non-standardized GDP

The table provides forecasting regressions of aggregate GDP on one period lagged components of GDP as in the main paper, except that the variables are not standardized. Newey-West corrected t-statistics are reported in brackets (automatic lag length selection). All data are annual and the sample period is from 1951 to 2010.

Forecasting regressions for real per capita growth rates:

$$\triangle GDP_t = \alpha + \beta \triangle Y_{j,t-1} + \epsilon_t$$

	α	$t(\alpha)$	β	$t(\beta)$	R^2
GDP	1.74	[4.43]	0.10	[0.75]	0.01
RES	1.88	[7.97]	0.08	[5.06]	0.20
DUR	1.67	[4.62]	0.07	[1.41]	0.04
NDU	0.87	[1.47]	0.54	[2.28]	0.09
EQS	1.97	[5.87]	-0.01	[-0.22]	0.00
BST	2.01	[7.30]	-0.07	[-1.88]	0.06

Low GDP Events and Stock Market Returns. We find in the main text, that mimicking portfolio factors of leading GDP components significantly load on the book-to-market factor (HML), and that mimicking portfolio factors of lagging GDP components significantly load on the momentum factor (WML). This pattern is visualized in Figure B.3 using the event time approach. We show MKT, SMB, HML, and WML in aggregate GDP and GDP components-event time. Low realizations of residential investment growth correspond to low HML returns. Low realizations of business structures growth correspond to low momentum returns.

Stock Market Crashes and GDP Components. Figure B.4 shows aggregate GDP and five GDP components during stock market "crashes". We distinguish between general stock market crashes measured by MKT, size crashes measured by SMB, value crashes measured by HML, and momentum crashes measured by WML.

The GDP-based Three-Factor Model in one Picture. Figure B.5 provides a close-up view for residential investment and business structures during value and momentum crashes. Value crashes coincide with low realizations of residential investment growth. Momentum crashes coincide with low realizations of business structures growth.

25 Size / Momentum Portfolios. In the main paper, we study 25 Fama-French portfolios sorted on size and book-to-market, and 10 momentum portfolios. As an alternative, we consider 25 portfolios sorted on size and momentum in Table B.5. We find that all GDP mimicking portfolio factors have a positive risk premium. Overall, the leading GDP components have a better model fit than the lagging GDP components. When we add business structures to residential investment (specification 8) the factor risk premium for BST is significant and the model fit is further improved.

Figure B.3: Low GDP Events and Stock Market Returns

The figure shows annual real returns for the market premium (MKT), the size premium (SMB) the book-to-market premium (HML) and the momentum premium (WML) in event time. The event is defined as years with the 30% lowest realizations of GDP growth and is labeled '0'. The years '-1', '-2, and '-3' refer to one, two, and three years before the event takes place. The years '1', '2, and '3' refer to one, two, and three years after the event takes place. The panels plot event time for aggregate GDP (GDP), residential investment (RES), durables (DUR), nondurables (NDU), equipment and software (EQS) and business structures (BST). The sample is from 1951 to 2010.

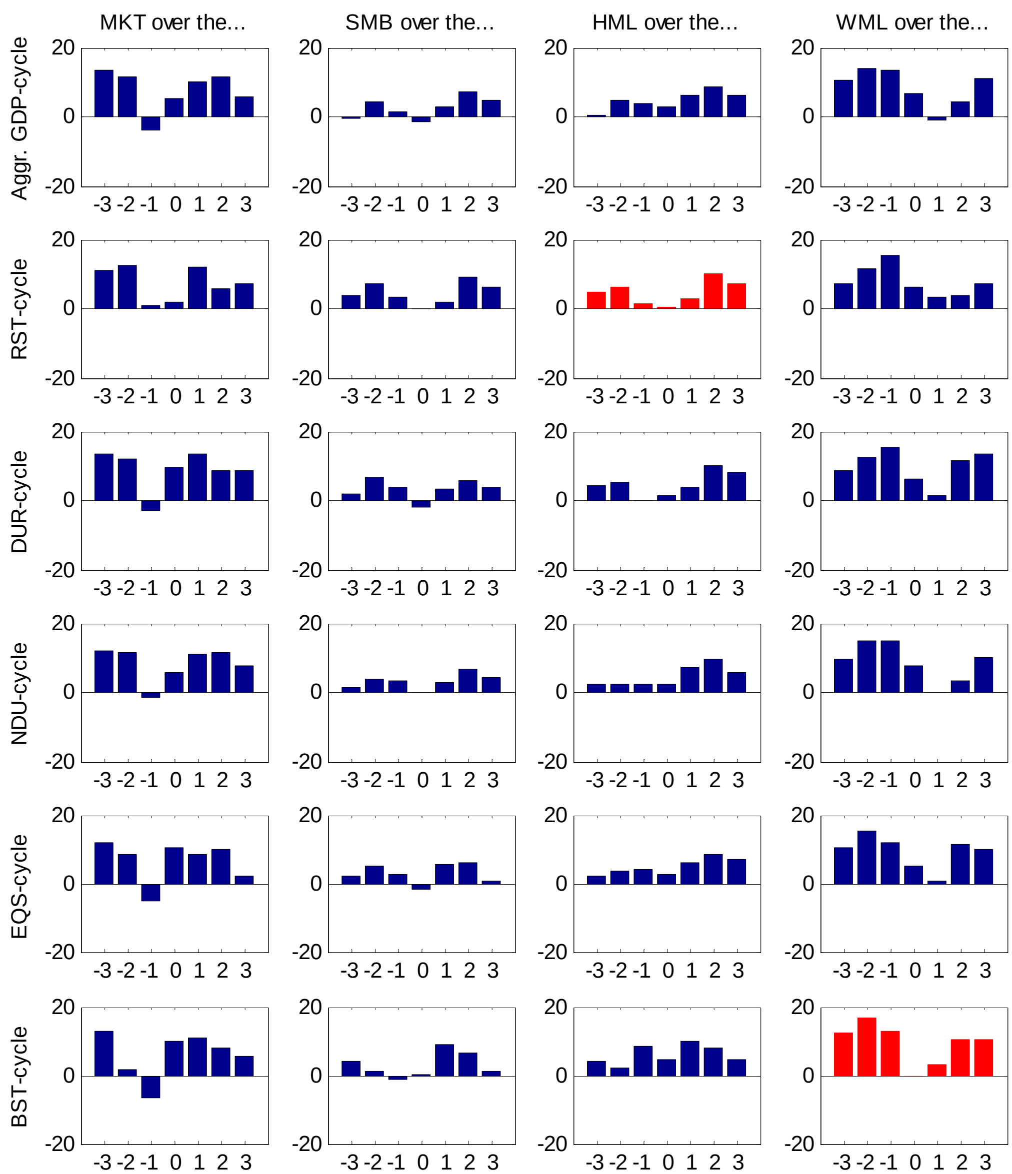


Figure B.4: Stock Market Crashes

The figure shows annual real per capita growth of aggregate GDP and five GDP components during stock market crashes. A stock market crash is defined as years with the 30% lowest realizations of returns and is labeled '0'. The years '-1', '-2, and '-3' refer to one, two, and three years before the event takes place. The years '1', '2, and '3' refer to one, two, and three years after the event takes place. The panels plot stock market crashes for the market premium (MKT), the size premium (SMB) the book-to-market premium (HML) and the momentum premium (WML). The sample is from 1951 to 2010.

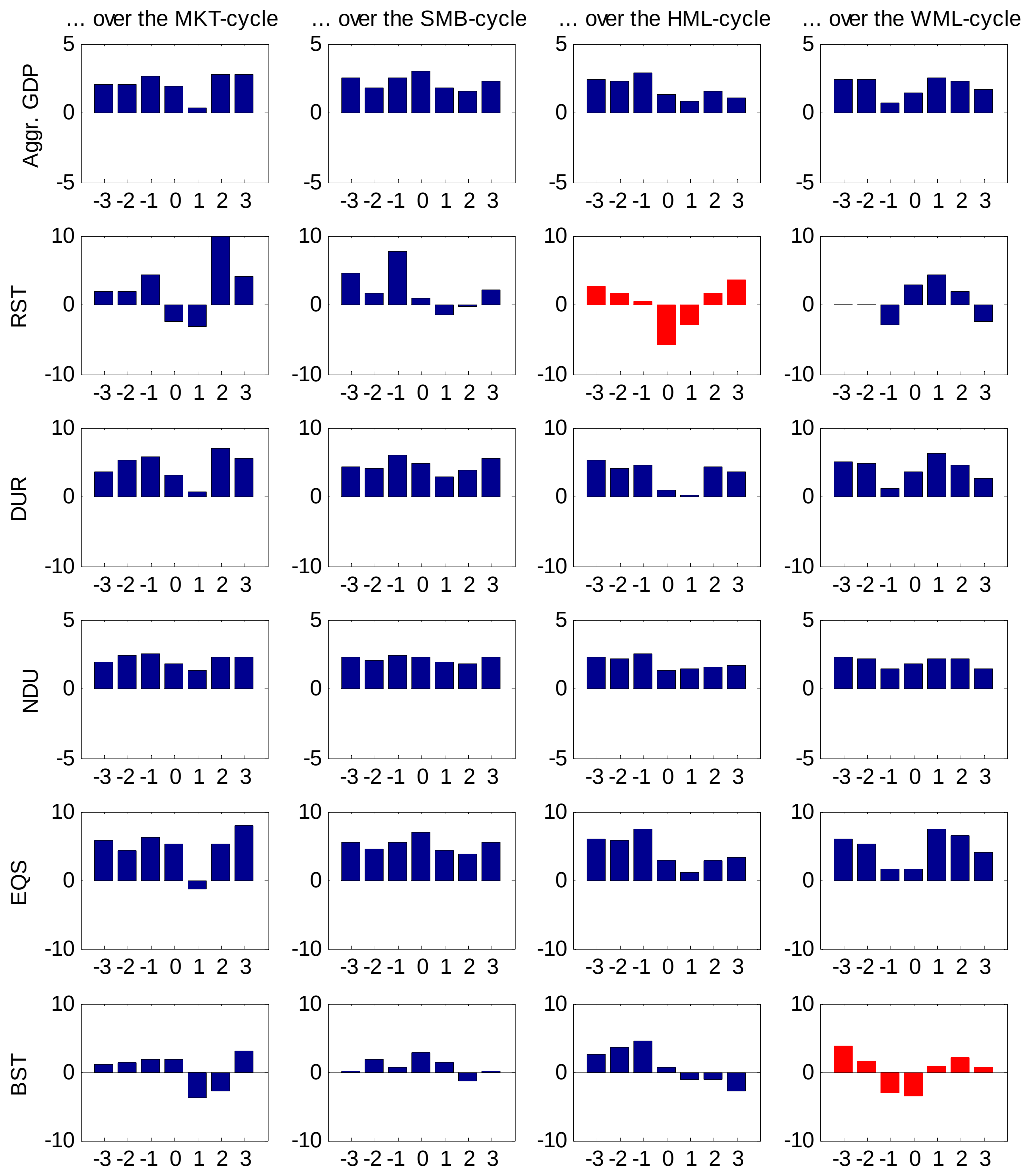
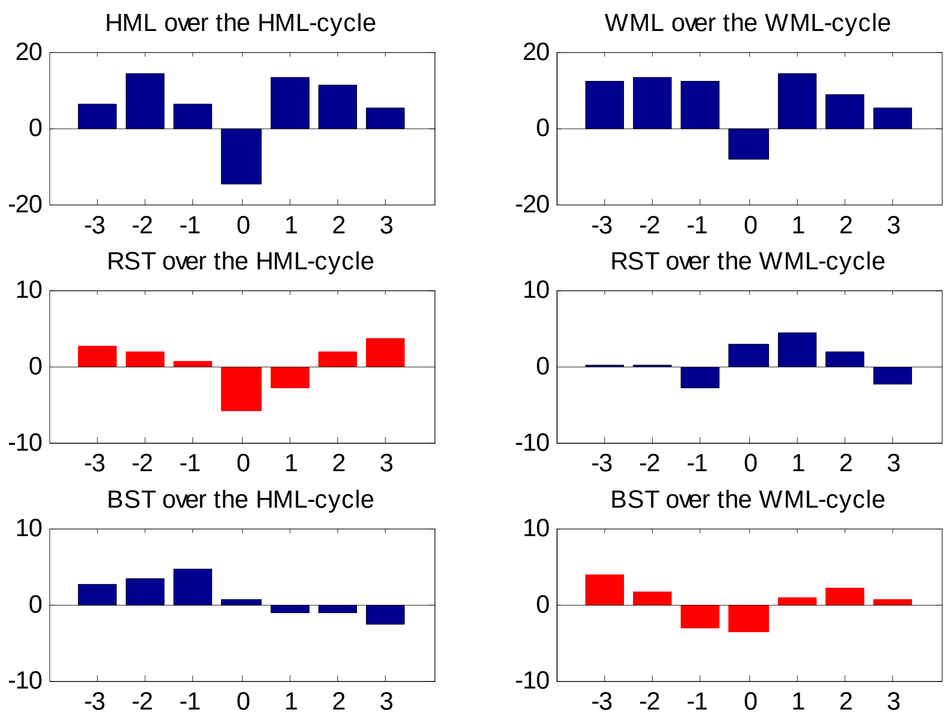


Figure B.5: Value and Momentum Crashes

The figure shows annual real per capita growth of residential investment (RST) and business structures (BST) during book-to-market (HML) and momentum (WML) crashes. A stock market crash event is defined as years with the 30% lowest realizations of returns and is labeled '0'. The years '-1', '-2, and '-3' refer to one, two, and three years before the event takes place. The years '1', '2, and '3' refer to one, two, and three years after the event takes place. The sample is from 1951 to 2010.



25 Fama-French or 10 Momentum Portfolios. How does the GDP-based three factor model perform on 25 Fama-French *or* 10 momentum portfolios? Table B.6 shows that the slope coefficient for BST is negative and insignificant if we include only 25 Fama-French portfolios. In contrast, the slope coefficient for RES is positive and significant. Results are vice versa if we include only 10 momentum portfolios.

This finding is not surprising. The main text shows that the growth rate of leading GDP component covaries only little with the momentum portfolios, and that the growth rate of lagging GDP component covaries only little with the 25 Fama-French portfolios. Thus, momentum portfolios are not informative for SDF loadings and factor prices of leading GDP components and vice versa.

Table B.5: GDP Mimicking Factors: Size / Momentum 25

The table reports GMM estimates of factor prices (λ) and SDF loadings (b) of the monthly GDP mimicking portfolio factors as described in the main paper. This table shows results for 25 double sorted size and momentum portfolios. T-statistics are reported in parentheses and are based on the Newey and West (1987) HAC covariance matrix with automatic lag length selection according to Andrews (1991). All data are monthly and the sample period is from January 1951 to December 2010.

	Const	SDF Loadings (b)			Factor Prices (λ)			R^{2OLS}_{GLS}	MAE	HJ(pv)
(1)	γ	MKT			MKT					
	1.56	-4.11			-0.77			0.10	0.28	0.36
	(5.49)	(-2.14)			(-2.14)			0.02		(0.00)
(2)	γ	MKT	GDP		MKT	GDP				
	0.27	5.34	5.05		0.58	0.90		0.68	0.16	0.33
	(1.07)	(2.63)	(3.48)		(1.81)	(3.80)		0.18		(0.00)
(3)	γ	MKT	RES		MKT	RES				
	-0.20	0.32	13.94		0.78	2.24		0.85	0.11	0.33
	(-0.44)	(0.12)	(3.59)		(1.62)	(3.97)		0.21		(0.00)
(4)	γ	MKT	DUR		MKT	DUR				
	-0.08	6.11	7.47		0.84	1.98		0.83	0.12	0.32
	(-0.22)	(2.43)	(3.56)		(2.10)	(4.29)		0.23		(0.00)
(5)	γ	MKT	NDU		MKT	NDU				
	0.03	4.50	8.39		0.76	1.05		0.76	0.14	0.33
	(0.11)	(2.24)	(3.54)		(2.22)	(4.79)		0.20		(0.00)
(6)	γ	MKT	EQS		MKT	EQS				
	0.44	4.54	4.23		0.42	1.17		0.72	0.15	0.33
	(1.69)	(2.27)	(3.49)		(1.30)	(3.88)		0.19		(0.00)
(7)	γ	MKT	BST		MKT	BST				
	0.51	4.23	3.90		0.37	1.07		0.61	0.18	0.34
	(2.23)	(2.39)	(3.34)		(1.27)	(3.47)		0.14		(0.00)
(8)	γ	MKT	RES	BST	MKT	RES	BST			
	-0.22	2.23	11.32	1.28	0.86	1.97	0.65	0.88	0.10	0.32
	(-0.52)	(1.20)	(3.14)	(1.00)	(2.03)	(3.49)	(1.96)	0.23		(0.00)

Table B.6: The GDP-based Model: Fama-French 25 or Momentum 10

The table reports GMM estimates of factor prices (λ) and SDF loadings (b) of the monthly GDP-based three factor model as described in the main paper. This table shows results for 25 Fama-French portfolios and 10 momentum portfolios, separately, with and without intercept. T-statistics are reported in parentheses and are based on the Newey and West (1987) HAC covariance matrix with automatic lag length selection according to Andrews (1991). All data are monthly and the sample period is from January 1951 to December 2010.

		γ	Factor Price (λ)			R^{2OLS}	γ	Factor Price (λ)			R^{2OLS}
		Fama-French 25					Momentum 10				
(1)		γ	MKT	RES	BST		γ	MKT	RES	BST	
	b	1.12	-6.03	4.86	-3.41	0.82	1.80	15.11	-40.14	15.49	0.94
		(3.30)	(-2.49)	(2.43)	(-1.55)		(1.54)	(1.58)	(-1.64)	(2.01)	
	λ		-0.51	0.35	-0.51	0.82		-0.91	-5.09	3.05	0.94
			(-1.41)	(0.81)	(-0.69)			(-0.83)	(-1.52)	(2.23)	
		Const	MKT	RES	BST		Const	MKT	RES	BST	
(2)	b	-	0.98	7.88	0.00	0.74	-	4.27	2.93	3.07	0.88
			(0.73)	(4.47)	(0.00)			(1.05)	(0.38)	(1.11)	
	λ		0.59	1.31	0.16	0.74		0.62	0.79	0.84	0.88
			(3.40)	(4.27)	(0.24)			(3.34)	(0.85)	(1.96)	

News Related to Future GDP

Vassalou (2003). We reconsider how important the lead and lag structure, or the timing, coherent in the GDP components actually is. Vassalou (2003) finds that a "news related to *future* GDP" mimicking portfolio can explain the cross-section of size/book-to-market portfolios. The weights for Vassalou's mimicking portfolio factor are found by regressing the aggregate GDP growth rate of the *next* year on asset returns of the *current* year.

Design. We investigate the importance of timing by constructing modified GDP mimicking portfolios which take the lead and lag structure into account. More precisely, we estimate mimicking portfolio weights as in Equation (B.7), except that we use the growth rate of aggregate GDP measured in $t+1$, residential investment in t, durables in $t+1$, nondurables in $t+1$, equipment and software in $t+1$, and business structures in $t+2$. This timing mainly eliminates the lead of residential investment and durables compared to all other variables at the annual frequency.

Results. Table B.7 reports the results for the Fama-French portfolios. Since we need future GDP variables (up to $t+2$) we use only asset return data from 1951 to 2008 for all seven specifications. After accounting for lead and lags, aggregate GDP as well as all five GDP components capture a substantial fraction of the cross-sectional variation in average returns.

We find that the specification with contemporaneous residential investment slightly outperforms future aggregate GDP (Vassalou (2003)). The constant for future aggregate GDP is about twice as large as for residential investment (8.5% vs 4.3%); only the SDF loading is significant (and positive) but not the risk factor price. Both models have approximately equal OLS R^2s

(0.82 vs 0.83). Note that after controlling for their lag, equipment and software as well as business structures show economically plausible positive point estimates for the SDF loadings and risk factor prices.

Table B.7: News Related to Future GDP: Fama-French 25

The table reports GMM estimates of a constant (γ), SDF loadings ($\mathbf{b}$), and factor prices ($\boldsymbol{\lambda}$) as in the main paper. The risk factors are the excess return on the market portfolio (MKT), and factor mimicking portfolios for contemporaneous period residential investment (RES), next period aggregate GDP (L^{-1}GDP), durables (L^{-1}DUR), nondurables (L^{-1}NDU), equipment and software (L^{-1}EQS), and two period ahead business structures (L^{-2}BST). Thus, the GDP mimicking portfolios control for lags to residential investment. The stock returns are 25 Fama-French portfolios sorted by size and book-to-market. All data are annual and the sample period is from 1951 to 2008/2010.

	Const	SDF Loadings ($\mathbf{b}$)		Factor Prices ($\boldsymbol{\lambda}$)		R^{2OLS}	MAE	HJ(pv)
(1)	γ	MKT	L^{-1}GDP	MKT	L^{-1}GDP			
	8.53	-6.88	6.39	-1.54	3.33	0.82	0.87	0.75
	(1.29)	(-1.75)	(2.60)	(-0.22)	(0.47)			(0.12)
(2)	γ	MKT	L^{0}RES	MKT	L^{0}RES			
	4.30	-0.64	1.10	2.99	34.44	0.83	0.88	0.73
	(0.70)	(-0.26)	(2.58)	(0.44)	(2.41)			(0.21)
(3)	γ	MKT	L^{-1}DUR	MKT	L^{-1}DUR			
	5.30	-2.73	3.57	1.88	10.72	0.82	0.92	0.74
	(0.91)	(-1.02)	(2.56)	(0.29)	(1.58)			(0.14)
(4)	γ	MKT	L^{-1}NDU	MKT	L^{-1}NDU			
	13.51	-7.58	4.60	-6.23	4.14	0.84	0.88	0.74
	(1.81)	(-1.77)	(2.44)	(-0.79)	(0.46)			(0.14)
(5)	γ	MKT	L^{-1}EQS	MKT	L^{-1}EQS			
	4.51	-4.93	7.25	3.15	8.53	0.75	1.07	0.75
	(0.72)	(-1.49)	(2.63)	(0.46)	(1.37)			(0.11)
(6)	γ	MKT	L^{-2}BST	MKT	L^{-2}BST			
	18.56	-8.04	3.66	-10.69	4.96	0.56	1.45	0.77
	(1.93)	(-2.14)	(2.58)	(-1.15)	(0.45)			(0.06)

Subsamples and Out-of-sample Estimation

Subsamples. We want to make sure that our cross-sectional regressions are not driven by a few occasional "outlier" or one particular event like the recent financial crisis. We estimate the mimicking portfolio weights using the full sample from 1951 to 2010, as in the main text. Following, we split our sample in two distinct subsamples from 1951 to 1980 and 1981 to 2010. Table B.8 and Table B.9 provide the cross-sectional slope coefficients for both subsamples. We find very similar results in both subsamples which are comparable to the full sample results in the main text.

Out-of-sample. We also study the stability of our mimicking portfolio factor weights in a challenging out-of-sample setup in Table B.10. We use the first 30 years of annual data (1951 to 1980) to calculate the monthly mimicking portfolio factors of the twelve months in the following year (1981). Following, we expand the estimation window by one year, and repeat for the next 12 months.

Thus, the mimicking portfolio factors use real time information only when running cross-sectional tests for the period 1981 to 2010 in Table B.10. We find that the resulting portfolio weights are very noisy when only a few observations are available (approximately less than 40 annual observations), but quickly stabilize for the further expanding samples. In this light, it is remarkable that the out-of-sample/expanding window results in Table B.10 are still comparable to the full sample results presented in the main text.

Innovations in GDP components

The innovation of a particular state variable should carry the relevant information for stock market returns. Given that GDP components lead and lag, we reconsider our results using the innovations of GDP components derived from a VAR. Campbell (1996) and Petkova (2006) propose a similar approach.

Estimation. First, we estimate the following first-order VAR:

$$\begin{bmatrix} \Delta RES_t \\ \Delta DUR_t \\ \Delta NDU_t \\ \Delta EQS_t \\ \Delta BST_t \end{bmatrix} = \mathbf{A}_0 + \mathbf{A}_1 \begin{bmatrix} \Delta RES_{t-1} \\ \Delta DUR_{t-1} \\ \Delta NDU_{t-1} \\ \Delta EQS_{t-1} \\ \Delta BST_{t-1} \end{bmatrix} + \boldsymbol{u}_t,$$

where $\boldsymbol{u}_t$ represents a vector of five innovations in GDP components. Second, we construct five mimicking portfolios for each element of $\boldsymbol{u}_t$, as is in the main paper for the simple growth rates. An alternative to the VAR approach would be to include lagged GDP components on the right hand side when estimating GDP mimicking portfolio factors, as in Lamont (2001) or Vassalou (2003). In unreported tests we find that this alternative procedure leads to very similar results as the VAR innovations.

Results. Table B.11 shows cross-sectional regression results for mimicking portfolio factors of GDP component VAR innovations. Overall, we find only little difference to the results based on parsimonious growth rates provided in the main text.

Quarterly GDP Data

We rely on annual GDP data in the main text. GDP estimates are also available at the quarterly frequency. However, there are several reasons to suspect that quarterly GDP data are more prone to measurement error than annual data - a fact which is often overlooked.

Intra-year estimates of GDP are based on less comprehensive monthly surveys. If no monthly survey is available, interpolation and other methods are used for intra-year estimates. Furthermore, intra-year estimates are subject to seasonality and thus seasonal adjustments which come naturally in different shades for the different GDP components (see e.g., Ferson and Harvey (1992) for a discussion of the seasonality issue and the impact of adjustments). For example, residential investment is more prone to seasonality than investment in equipment and software.

Table B.8: Subsample: 1951 to 1980

The table reports GMM estimates of a constant (γ) and factor prices ($\boldsymbol{\lambda}$) as in the main paper, except that we use only use the first 30 years for the cross-sectional regressions (January 1951 to December 1980). T-statistics are reported in parentheses and are based on the Newey and West (1987) HAC covariance matrix with automatic lag length selection according to Andrews (1991).

	γ	Factor Price ($\boldsymbol{\lambda}$)		R^{2OLS}	γ	Factor Price ($\boldsymbol{\lambda}$)		R^{2OLS}	γ	Factor Price ($\boldsymbol{\lambda}$)			R^{2OLS}
	Fama-French 25				Momentum 10				Fama-French 25 + Mom. 10				
(1)	γ	MKT	RES		γ	MKT	RES		γ	MKT	RES		
	0.39	0.29	0.81	0.65	0.62	0.04	3.04	0.89	0.33	0.28	1.18		0.55
	(0.78)	(0.50)	(2.08)		(1.01)	(0.06)	(2.69)		(0.71)	(0.50)	(3.15)		
(2)	γ	MKT	DUR		γ	MKT	DUR		γ	MKT	DUR		
	0.09	0.68	1.31	0.45	0.10	0.57	2.56	0.90	-0.25	0.98	2.10		0.68
	(0.16)	(1.12)	(2.87)		(0.20)	(1.01)	(4.02)		(-0.47)	(1.60)	(5.53)		
(3)	γ	MKT	NDU		γ	MKT	NDU		γ	MKT	NDU		
	-0.16	0.89	0.72	0.21	0.19	0.47	1.12	0.92	-0.53	1.21	1.13		0.68
	(-0.29)	(1.39)	(1.79)		(0.42)	(0.90)	(4.93)		(-0.99)	(1.93)	(5.55)		
(4)	γ	MKT	EQS		γ	MKT	EQS		γ	MKT	EQS		
	1.27	-0.61	-1.16	0.19	0.05	0.61	1.33	0.92	-0.19	0.98	0.84		0.41
	(1.93)	(-0.89)	(-1.40)		(0.10)	(1.15)	(3.83)		(-0.39)	(1.68)	(2.38)		
(5)	γ	MKT	BST		γ	MKT	BST		γ	MKT	BST		
	1.80	-1.10	-1.23	0.48	0.10	0.55	1.42	0.93	-0.26	1.00	0.59		0.30
	(2.83)	(-1.72)	(-1.90)		(0.23)	(1.05)	(3.68)		(-0.53)	(1.78)	(1.56)		
(6)									γ	MKT	RES	BST	
									-0.25	0.91	1.49	0.68	0.75
									(-0.52)	(1.59)	(3.78)	(1.82)	

Table B.9: Subsample: 1981 to 2010

The table reports GMM estimates of a constant (γ) and factor prices ($\boldsymbol{\lambda}$) as in the main paper, except that we use only use the last 30 years for the cross-sectional regressions (January 1981 to December 2009). T-statistics are reported in parentheses and are based on the Newey and West (1987) HAC covariance matrix with automatic lag length selection according to Andrews (1991).

	γ	Factor Price ($\boldsymbol{\lambda}$)		R^{2OLS}	γ	Factor Price ($\boldsymbol{\lambda}$)		R^{2OLS}	γ	Factor Price ($\boldsymbol{\lambda}$)			R^{2OLS}
	Fama-French 25				Momentum 10				Fama-French 25 + Mom. 10				
(1)	γ	MKT	RES		γ	MKT	RES		γ	MKT	RES		
	0.32	0.28	1.32	0.67	0.51	0.05	0.92	0.67	0.34	0.25	1.23		0.68
	(0.44)	(0.34)	(2.17)		(0.84)	(0.08)	(1.20)		(0.53)	(0.35)	(2.55)		
(2)	γ	MKT	DUR		γ	MKT	DUR		γ	MKT	DUR		
	0.25	0.46	3.00	0.71	0.69	-0.10	0.94	0.70	0.89	-0.20	1.34		0.57
	(0.34)	(0.56)	(3.74)		(1.43)	(-0.19)	(1.56)		(1.43)	(-0.30)	(2.52)		
(3)	γ	MKT	NDU		γ	MKT	NDU		γ	MKT	NDU		
	1.57	-0.83	0.40	0.39	0.61	-0.03	0.54	0.69	1.00	-0.30	0.59		0.45
	(3.00)	(-1.63)	(0.59)		(1.21)	(-0.05)	(1.58)		(1.57)	(-0.46)	(1.93)		
(4)	γ	MKT	EQS		γ	MKT	EQS		γ	MKT	EQS		
	1.75	-0.97	1.46	0.39	0.83	-0.23	0.83	0.71	1.45	-0.73	1.03		0.43
	(3.10)	(-1.59)	(1.65)		(1.98)	(-0.44)	(1.70)		(3.01)	(-1.33)	(2.20)		
(5)	γ	MKT	BST		γ	MKT	BST		γ	MKT	BST		
	2.45	-1.84	-0.83	0.59	0.75	-0.15	0.79	0.71	1.58	-0.87	0.80		0.36
	(3.97)	(-2.76)	(-0.65)		(1.72)	(-0.28)	(1.67)		(3.12)	(-1.55)	(1.82)		
(6)									γ	MKT	RES	BST	
									0.38	0.19	1.25	0.39	0.70
									(0.55)	(0.26)	(2.56)	(0.81)	

Table B.10: Out-of-sample: Expanding Window

The table reports GMM estimates of a constant (γ) and factor prices ($\boldsymbol{\lambda}$) as in the main paper, except that we use a rolling window to estimate the GDP mimicking portfolio weights and run cross-sectional regressions "out-of-sample". We use the first 30 years of annual data (1951 to 1980) to calculate the monthly mimicking portfolio factors of the twelve months in the following year (1981). We expand the estimation window by one year, and repeat for the next 12 months. The table reports cross-sectional regressions of the out-of-sample generated monthly factor mimicking portfolios from January 1981 to December 2010. T-statistics are reported in parentheses and are based on the Newey and West (1987) HAC covariance matrix with automatic lag length selection according to Andrews (1991).

	γ	Factor Price ($\boldsymbol{\lambda}$)		R^{2OLS}	γ	Factor Price ($\boldsymbol{\lambda}$)		R^{2OLS}	γ	Factor Price ($\boldsymbol{\lambda}$)			R^{2OLS}
	FF25				Momentum 10				FF25 + M10				
(1)	γ	MKT	RES		γ	MKT	RES		γ	MKT	RES		
	0.68	-0.07	0.29	0.60	1.28	-0.63	-0.28	0.75	1.62	-0.95	0.00		0.35
	(0.97)	(-0.09)	(1.99)		(2.53)	(-1.07)	(-1.32)		(3.53)	(-1.73)	(0.04)		
(2)	γ	MKT	DUR		γ	MKT	DUR		γ	MKT	DUR		
	0.60	0.04	0.33	0.76	0.69	-0.11	0.16	0.68	0.75	-0.10	0.23		0.67
	(0.74)	(0.05)	(3.03)		(1.38)	(-0.20)	(1.57)		(1.17)	(-0.14)	(3.17)		
(3)	γ	MKT	NDU		γ	MKT	NDU		γ	MKT	NDU		
	0.37	0.22	0.03	0.68	0.34	0.16	0.04	0.53	0.36	0.20	0.03		0.64
	(0.49)	(0.26)	(2.59)		(0.31)	(0.15)	(0.92)		(0.55)	(0.27)	(3.08)		
(4)	γ	MKT	EQS		γ	MKT	EQS		γ	MKT	EQS		
	1.73	-0.99	0.80	0.67	0.92	-0.32	0.21	0.71	1.43	-0.72	0.33		0.52
	(2.12)	(-1.18)	(3.52)		(2.27)	(-0.62)	(1.51)		(3.08)	(-1.35)	(2.55)		
(5)	γ	MKT	BST		γ	MKT	BST		γ	MKT	BST		
	1.72	-0.96	0.31	0.40	0.84	-0.24	0.14	0.71	1.44	-0.74	0.18		0.44
	(2.95)	(-1.57)	(1.54)		(2.07)	(-0.47)	(1.65)		(3.12)	(-1.40)	(2.10)		
(6)									γ	MKT	RES	BST	
									0.18	0.41	0.28	0.17	0.68
									(0.28)	(0.57)	(1.97)	(2.22)	

Table B.11: Innovations in GDP Components

The table reports GMM estimates of a constant (γ) and factor prices (λ) as in the main paper, except that we use monthly GDP mimicking portfolios based on the innovations of a VAR system containing the five GDP components. T-statistics are reported in parentheses and are based on the Newey and West (1987) HAC covariance matrix with automatic lag length selection according to Andrews (1991). The sample period is from January 1951 to December 2010.

	γ	Factor Price (λ)		R^{2OLS}	γ	Factor Price (λ)		R^{2OLS}	γ	Factor Price (λ)			R^{2OLS}
	FF25				Momentum 10				FF25 + M10				
(1)	γ	MKT	RES		γ	MKT	RES		γ	MKT	RES		
	0.50	0.12	0.98	0.83	0.17	0.42	1.31	0.88	0.37	0.23	1.12		0.87
	(1.13)	(0.24)	(2.64)		(0.49)	(1.12)	(3.25)		(0.92)	(0.50)	(3.46)		
(2)	γ	MKT	DUR		γ	MKT	DUR		γ	MKT	DUR		
	0.67	0.12	1.97	0.58	0.23	0.41	1.55	0.90	0.52	0.22	1.67		0.71
	(1.39)	(0.22)	(4.92)		(0.64)	(1.03)	(3.54)		(1.24)	(0.47)	(4.27)		
(3)	γ	MKT	NDU		γ	MKT	NDU		γ	MKT	NDU		
	-0.42	1.08	1.21	0.51	-0.22	0.80	1.18	0.88	-0.51	1.14	1.32		0.71
	(-0.83)	(1.88)	(3.08)		(-0.52)	(1.83)	(3.49)		(-1.03)	(2.12)	(4.40)		
(4)	γ	MKT	EQS		γ	MKT	EQS		γ	MKT	EQS		
	1.43	-0.61	1.11	0.16	0.61	0.05	1.62	0.91	1.21	-0.42	1.73		0.48
	(4.04)	(-1.47)	(2.35)		(1.87)	(0.14)	(3.57)		(3.54)	(-1.06)	(4.05)		
(5)	γ	MKT	BST		γ	MKT	BST		γ	MKT	BST		
	0.31	0.37	1.66	0.32	0.35	0.27	1.14	0.90	0.44	0.25	1.28		0.64
	(0.68)	(0.70)	(2.62)		(1.17)	(0.75)	(3.46)		(1.19)	(0.61)	(4.82)		
(6)									γ	MKT	RES	BST	
									0.34	0.27	1.10	0.98	0.87
									(0.82)	(0.61)	(3.41)	(3.45)	

Estimation. To reduce these potential problems, we apply a two quarters moving average for all five GDP components at the quarterly frequency:

$$\Delta\tilde{Y}_{j,t} = \frac{Y_{j,t} + Y_{j,t-1}}{Y_{j,t-1} + Y_{j,t-2}} - 1.$$

This averaging adjustment is also used by Jagannathan and Wang (1996) for labor income growth as a risk factor. Following, we calculate GDP mimicking portfolio weights based on $\Delta\tilde{Y}_{j,t}$ and use these weights with monthly stock market data.

Results. Table B.12 shows results for the GDP mimicking portfolio factors based on quarterly data. Overall, the mimicking portfolio weights and factor loadings are very similar as in the main text. Leading GDP components load on HML, lagging GDP components load on WML.

One exception is observed for DUR. Using quarterly data, the weights on the mid/big and high/big portfolio are extreme, resulting in a large monthly standard deviation of 10.2%. In Table B.13 we find similar cross-sectional results for the monthly GDP mimicking portfolios based on quarterly GDP data as we find for monthly GDP mimicking portfolios based on annual GDP data in the main text. Only results for the DUR factor (specification 2) are considerably inferior.

Table B.12: GDP Mimicking Portfolios Based on Quarterly Data

Panel A reports estimates of $j = 1, ..., 6$ mimicking portfolio weights for aggregate GDP (GDP), residential investment (RES), durables (DUR), nondurables (NDU), equipment and software (EQS) and business structures (BST). We regress the real per capita growth rate of quarterly aggregate GDP and each of the quarterly GDP components ($\triangle\tilde{Y}_{j,t}$) on six annual Fama-French portfolios, $FF6_t$, sorted by size and book-to-market (low/small, mid/small, high/small, low/big, mid/big, high/big) and the momentum factor, WML_t:

$$\triangle\tilde{Y}_{j,t} = a_j + \mathbf{p}_j' \left[\mathbf{FF6}_t',\ WML_t\right]' + \epsilon_{j,t}.$$

The resulting weights $\hat{\mathbf{p}}_j$ are normalized, such that their sum is one: $\hat{\mathbf{w}}_j = \hat{\mathbf{p}}_j \left(\mathbf{1}'\hat{\mathbf{p}}_{,j}\right)^{-1}$. Monthly GDP mimicking portfolios are calculated as: $MPm_{j,t} = \hat{\mathbf{w}}_j' \left[\mathbf{FF6m}_t',\ WMLm_t\right]'$, where $\mathbf{FF6m}_t$, and $WMLm_t$ are monthly measured returns of the six Fama-French portfolios and the momentum factor. Panel B reports the mean (%), standard deviation (%), Sharpe ratio and the factor exposures (betas) of the six monthly GDP mimicking portfolios:

$$MPm_{j,t} = \alpha_j + \beta_{MKT,j}MKT_t + \beta_{SMB,j}SMB_t + \beta_{HML,j}HML_t + \beta_{WML,j}WML_t,$$

where MKT is the market excess return, SMB is the small minus big factor, HML is the high minus low factor, and WML is the winner minus loser (momentum) factor. The sample period is from January 1951 to December 2010.

	GDP	RES	DUR	NDU	EQS	BST
Panel A:	Normalized GDP mimicking portfolio weights					
low/small	-0.76	-0.74	-2.11	-0.24	-1.02	0.38
mid/small	0.23	0.54	-0.26	0.40	-2.86	-0.30
high/small	0.42	0.53	2.24	-0.24	3.16	-1.62
low/big	0.42	0.27	1.35	-0.26	0.42	-0.57
mid/big	-1.07	-0.49	-4.22	-0.36	0.39	-0.75
high/big	1.07	0.40	3.43	1.07	-0.20	2.46
momentum	0.69	0.50	0.58	0.63	1.11	1.40
Panel B:	Statistics of monthly mimicking portfolio factors					
Mean (%)	1.13	1.15	2.14	0.91	1.32	0.42
SD (%)	3.84	3.14	10.24	3.24	7.02	7.37
Sharpe ratio	0.30	0.37	0.21	0.28	0.19	0.06
α	0.00	0.00	0.00	-0.00	0.00	-0.00
s.e.	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
β_{MKT}	0.35	0.47	0.63	0.41	-0.07	-0.19
s.e.	(0.02)	(0.01)	(0.09)	(0.01)	(0.04)	(0.04)
β_{SMB}	-0.15	0.15	-0.08	-0.04	-0.74	-1.06
s.e.	(0.03)	(0.02)	(0.13)	(0.02)	(0.05)	(0.06)
β_{HML}	0.93	0.82	2.94	0.81	1.33	0.46
s.e.	(0.03)	(0.02)	(0.14)	(0.02)	(0.07)	(0.06)
β_{WML}	0.69	0.51	0.59	0.61	1.15	1.34
s.e.	(0.03)	(0.02)	(0.09)	(0.01)	(0.04)	(0.04)

Table B.13: GDP Mimicking Portfolio Factors Based on Quarterly Data

The table reports GMM estimates of a constant (γ) and factor prices ($\boldsymbol{\lambda}$) as in the main paper, except that we use monthly GDP mimicking portfolios based on quarterly GDP data. T-statistics are reported in parentheses and are based on the Newey and West (1987) HAC covariance matrix with automatic lag length selection according to Andrews (1991). The sample period is from January 1951 to December 2010.

	γ	Factor Price ($\boldsymbol{\lambda}$)		R^{2OLS}	γ	Factor Price ($\boldsymbol{\lambda}$)		R^{2OLS}	γ	Factor Price ($\boldsymbol{\lambda}$)			R^{2OLS}
	FF25				Momentum 10				FF25 + M10				
(1)	γ	MKT	RES		γ	MKT	RES		γ	MKT	RES		
	0.19	0.41	0.98	0.74	0.12	0.46	1.08	0.88	0.14	0.45	1.04		0.84
	(0.42)	(0.79)	(2.75)		(0.37)	(1.26)	(3.59)		(0.34)	(0.99)	(3.87)		
(2)	γ	MKT	DUR		γ	MKT	DUR		γ	MKT	DUR		
	0.49	0.21	1.87	0.44	-1.36	1.74	15.96	0.41	0.63	0.02	2.03		0.29
	(1.16)	(0.43)	(3.08)		(-0.58)	(0.84)	(1.05)		(1.70)	(0.06)	(3.26)		
(3)	γ	MKT	NDU		γ	MKT	NDU		γ	MKT	NDU		
	0.40	0.29	0.80	0.33	0.21	0.39	0.96	0.88	0.16	0.49	1.01		0.65
	(0.98)	(0.62)	(2.25)		(0.66)	(1.09)	(3.81)		(0.43)	(1.15)	(4.33)		
(4)	γ	MKT	EQS		γ	MKT	EQS		γ	MKT	EQS		
	1.07	-0.27	0.49	0.14	-0.09	0.71	1.73	0.86	0.07	0.68	1.28		0.37
	(3.72)	(-0.86)	(1.00)		(-0.20)	(1.47)	(3.61)		(0.16)	(1.44)	(3.99)		
(5)	γ	MKT	BST		γ	MKT	BST		γ	MKT	BST		
	2.15	-1.49	-1.18	0.61	0.31	0.33	1.50	0.89	1.06	-0.30	0.76		0.17
	(4.64)	(-3.17)	(-1.80)		(0.95)	(0.86)	(3.45)		(3.03)	(-0.77)	(2.23)		
(6)									γ	MKT	RES	BST	
									0.19	0.39	1.01	0.72	0.84
									(0.47)	(0.89)	(3.86)	(2.02)	

C Foreign Exchange Investment Styles

C.1 Data Archive

Currency Data

Data Sources. Spot and one-month forward exchange rate quotes are obtained from WMR / Reuters (WMR) and Barclays Bank International (BBI) available via Thomsen Reuters Datastream. Each month, we select the last trading day. We use WMR dollar quotes as soon as they are available starting in 12/1996. For the period from 12/1983 to 11/1996 we base our data set on BBI dollar quotes. BBI data for AUD, NZD, NOK, SEK, and DKK are only reported beginning in 12/1984, there are no BBI data available for ESP.

We apply the approach of Burnside (2011), using exchange rate quotes from WMR against the British pound swapped to the USD for the period from 01/1976 to 11/1983 (and 11/1984 for the currencies listed above), and beyond if there are no BBI quotes available. There are no "WMR pound" forward quotes for AUD and NZD before 12/1984 (the beginning of the BBI data) and the time series for the JPY starts in 06/1978. Some of the non-G10 currencies are not available for the full sample period.[75] The time series of the Deutschmark is spliced with the introduction of the Euro in 01/1999.

Table C.1 provides an overview of the compiled dataset including the Datastream mnemonics and dates. We compile the dataset this way since it guarantees the availability of mid, bid, and ask quotes for spot and forward rates for the listed currencies and for the covered time period.

Data Screens. We apply three careful screens to adjust our dataset of exchange rate quotes:

- First, following Lustig, Roussanov, and Verdelhan (2011) ("based on large failures of covered interest rate parity") and Menkhoff, Sarno, Schmeling, and Schrimpf (2012b) ("tradeability issues") we delete the following observations from our sample: South Africa 1985 2001; Hungary 2000; Indonesia 1999-2007; Philipines 1999, Taiwan 1999; Thailand 1999; Turkey 2000 2001. Note that G10 currencies are never deleted from the sample.

- Second, we find for a few WMR pound observations that some bid quotes are larger than ask quotes. In these rare cases, we set the corresponding quotes equal to the largest bid-ask spread in the remaining G10 currency quotes. In a few cases, bid-ask spreads in the WMR pound data are very small. To correct for this, we replace all bid-ask spreads which are smaller than 0.1% p.a. with the largest bid-ask spread of the G10 currencies to be conservative.

- Third, in the WMR pound and BBI data, there are a few occasional blips in the forward premium, $F_t/S_t - 1$. That is, the forward premium deviates from the "synthetic forward premium" implied by covered interest rate parity for one particular observation but not for the observations one period before and after. We identify such outliers as an increase or decrease of the forward premium by (more than) 3% from one period to the other

[75]G10 are the currencies of Australia, Canada, Germany, Japan, New Zealand, Norway, Sweden, Switzerland, and the United Kingdom against the USD.

and a subsequent reversal of (more than) 3% in the next period.[76] In these cases, we linearly interpolate the forward premium from both observations surrounding the outlier and calculate a corrected forward rate from the corresponding spot rate. We keep the bid-ask spread of the reported forward data and calculate corrected forward bid and ask quotes.

Table C.1: Datastream Mnemonics

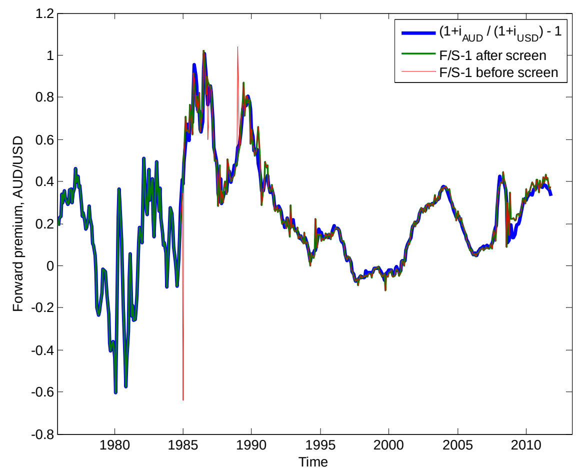
	WMR Dollar			BBI Dollar			WMR Pound		
	Spot	Forward	Date	Spot	Forward	Date	Spot	Forward	Date
Australia	AUSTDO$	USAUD1F	12/1996-12/2011	BBAUDSP	BBAUD1F	12/1984-11/1996	AUSTDOL	-	01/1976-11/1983
Canada	CNDOLL$	USCAD1F	12/1996-12/2011	BBCADSP	BBCAD1F	12/1984-11/1996	CNDOLLR	CNDOL1F	01/1976-11/1983
Germany	DMARKE$	USDEM1F	12/1996-11/1998	BBDEMSP	BBDEM1F	12/1983-11/1996	DMARKER	DMARK1F	01/1976-11/1983
Japan	JAPAYE$	USJPY1F	12/1996-12/2011	BBJPSP	BBJP1F	12/1983-11/1996	JAPAYEN	JAPYN1F	06/1978-11/1983
New Zeal.	NZDOLL$	USNZD1F	12/1996-12/2011	BBNZDSP	BBNZD1F	12/1984-11/199	NZDOLLR	-	01/1976-11/1983
Norway	NORKRO$	USNOK1F	12/1996-12/2011	BBNOKSP	BBNOK1F	12/1984-11/1996	NORKRON	NORKN1F	01/1976-11/1983
Sweden	SWEKRO$	USSEK1F	12/1996-12/2011	BBSEKSP	BBSEK1F	12/1984-11/1996	SWEKRON	SWEDK1F	01/1976-11/1983
Switzerl.	SWISSF$	USCHF1F	12/1996-12/2011	BBCHFSP	BBCHF1F	12/1983-11/1996	SWISSFR	SWISF1F	01/1976-11/1983
U.K.	UKDOLLR	USGBP1F	12/1996-12/2011	BBGBPSP	BBGBP1F	12/1983-11/1996	USDOLLR	USDOL1F	01/1976-11/1983
Euro zone	EUDOLLR	USEUR1F	12/1998-12/2011	-	-		-	-	
France	FRENFR$	USFRF1F	12/1996-12/1998	BBFRSP	BBFR1F	12/1983-11/1996	FRENFRA	FRENF1F	01/1976-12/1983
Italy	ITALIR$	USITL1F	12/1996-12/1998	BBITLSP	BBITL1F	12/1983-11/1996	ITALIRE	ITALY1F	01/1976-12/1983
Spain	SPANPE$	USESP1F	12/1996-12/1998	-	-		SPANPES	SPANP1F	01/1976-12/1996
Denmark	DANISH$	USDKK1F	12/1996-12/2011	BBDKSP	BBDK1F	12/1984-11/1996	DANISHK	DANIS1F	01/1976-12/1983
South A.	COMRAN$	USZAR1F	12/1996-12/2011	BBZARSP	BBZAR1F	12/1983-11/1996	-	-	
Brazil	BRACRU$	USBRL1F	03/2004-12/2011	-	-		-	-	
Hungary	HUNFOR$	USHUF1F	10/1997-12/2011	-	-		-	-	
Iceland	ICEKRO$	USISK1F	03/2004-12/2011	-	-		-	-	
India	INDRUP$	USINR1F	10/1997-12/2011	-	-		-	-	
Indonesia	INDORU$	USIDR1F	12/1996-12/2011	-	-		-	-	
Israel	ISRSHE$	USILS1F	03/2004-12/2011	-	-		-	-	
Mexico	MEXPES$	USMXN1F	12/1996-12/2011	-	-		-	-	
Philippines	PHILPE$	USPHP1F	12/1996-12/2011	-	-		-	-	
Poland	POLZLO$	USPHP1F	02/2002-12/2011	-	-		-	-	
Russia	CISRUB$	USRUB1F	03/2004-12/2011	-	-		-	-	
S. Korea	KORSWO$	USKRW1F	02/2002-12/2011	-	-		-	-	
Taiwan	TAIWDO$	USTWD1F	12/1996-12/2011	-	-		-	-	
Thailand	THABAH$	USTHB1F	12/1996-12/2011	-	-		-	-	
Turkey	TURKLI$	USTRY1F	12/1996-12/2011	-	-		-	-	
Ukraine	UKRAHY$	USUAH1F	03/2004-12/2011	-	-		-	-	

[76] For example, for AUD, we identify three such “blips”.

AUD, NZD and JPY forwards. There are no AUD and NZD forward quotes (and bid-ask spreads) available for the period 01/1976 to 11/1983. The same is true for JPY for the period 01/1976 to 06/1978. Spot quotes are available in the WMR pounds data set. Thus, we calculate synthetic forward rates for these three cases to have a complete sample of G10 currencies.[77] To have bid and ask quotes, we set the bid ask spread of the spot and forward rates to the largest bid-ask spread in the remaining G10 currency quotes to be conservative.

As an illustrative example, Figure C.1 shows the forward premium for AUD/USD implied by covered interest rate parity, before our data screens, and after our data screens based on forward and spot rates.

Figure C.1: AUD/USD Forward Premium



Transaction Costs

Our portfolios take into account transaction costs due to portfolio re-balancing. Each period, we discriminate between the fraction of individual currency portfolio weights $(w^z_{j;t})$ that changes and correct the corresponding individual currency return $RX_{j;t+1}$. In particular, for the amount of dollar long in currency j which remains constant between periods $t-1$, t, and $t+1$, we calculate the return as $\left(S^M_{j;t+1} - F^M_{j;t}\right)/S^M_{j;t}$, where M denotes mid exchange rate quotes. For the amount of dollar long in currency j which is increased from period $t-1$ to t, and which remains in the portfolio in period $t+1$, we calculate the return as $\left(S^M_{j;t+1} - F^A_{j;t}\right)/S^A_{j;t}$, where A denotes ask exchange rate quotes. For the amount of dollar long in currency j which is decreased from period t to $t+1$, and which is already in the portfolio in period $t-1$, we calculate the return as $\left(S^B_{j;t+1} - F^M_{j;t}\right)/S^M_{j;t}$, where B denotes bid exchange rate quotes. Finally, for the amount of dollar long in currency j which is increased from period $t-1$ to t, and which is decreased in period

[77] More specifically, we assume that covered interest rate parity holds and use short-term interest rates of these three countries and the U.S. from the IMF's International Financial Statistics (IFS) to calculate synthetic forwards. Using the IFS interest rates to calculate forwards is also followed by Campbell, Sunderam, and Viceira (2010).

$t+1$, we calculate the return as $\left(S^{B}_{j;t+1} - F^{A}_{j;t}\right)/S^{A}_{j;t}$. Short positions are calculated vice versa.

Comparison of FX Style Portfolios

We compare our FX investment style portfolios with similar FX portfolios studied by Lustig, Roussanov, and Verdelhan (2011), Burnside, Eichenbaum, and Rebelo (2011), and Asness, Moskowitz, and Pedersen (2012). The portfolios used for the comparisons are available at the following authors' websites (December 2012):

- Lustig, Roussanov, and Verdelhan (2011):
  http://web.mit.edu/adrienv/www/Data.html

- Burnside, Eichenbaum, and Rebelo (2011):
  http://people.duke.edu/~acb8/Craig_Burnsides_Web_Page/...
  Currency_Speculation_Data.html

- Asness, Moskowitz, and Pedersen (2012):
  http://faculty.chicagobooth.edu/tobias.moskowitz/research/chronology.html

Our three FX strategies are described in detail in the main text, and are based on up to 30 currencies ("all currencies") or a smaller subset of G10 currencies. Lustig, Roussanov, and Verdelhan (2011) provide returns for currency portfolios based on up to 35 currencies and a smaller subset of 15 developed market currencies. The portfolio returns of Burnside, Eichenbaum, and Rebelo (2011) are based on up to 20 currencies. Finally, Asness, Moskowitz, and Pedersen (2012) use the G10 currencies.

The Lustig, Roussanov, and Verdelhan (2011) and Burnside, Eichenbaum, and Rebelo (2011) portfolios are equally weighted across currencies. The quantiles chosen to construct long and short positions, however, differ. Lustig, Roussanov, and Verdelhan (2011) choose 3/15 quantiles, and Burnside, Eichenbaum, and Rebelo (2011) construct portfolios which are long in all currencies with a specific sign of the instrument (e.g. a positive interest rate differential) and short vice versa. Asness, Moskowitz, and Pedersen (2012) propose the cross-sectional rank weighting scheme which we follow and describe in the main paper.

Tables OA.2, OA.3 and OA.4 report descriptive statistics of the FX styles. These tables also show a comparison of our FX style portfolios to similar FX portfolios by other authors. Since the latter are only available over shorter time spans, the samples in this analysis run from **11/1983 - 10/2011** and **02/1981 - 06/2010**.

Table C.2: Comparison of FX Style Portfolios: Carry Trade

C refers to our own carry trade portfolio which is based on up to 30 currencies ("all currencies") or the G10 currencies. LRV is the carry trade portfolio as in Lustig, Roussanov, and Verdelhan (2011), where "all" refers to their portfolio with all currencies, and "dev" to their portfolio with developed currencies (up to 15). BER is the carry trade portfolio as in Burnside, Eichenbaum, and Rebelo (2011), which is based on up to 20 currencies. The mean is annualized by multiplying by 12, the standard deviation and the Sharpe ratio (SR) are annualized by multiplying by $\sqrt{12}$. The sample period is from 11/1983 to 10/2010.

	Carry trade, sample: 11/1983 - 10/2011							
	Mean	Std	Skew	Ac1	SR	Correlation		
LRV C all	8.32	9.04	-0.64	0.11	0.92	1.00		
LRV C dev	5.92	9.78	-0.88	0.10	0.61	0.65	1.00	
BER C	4.43	5.53	-0.68	0.09	0.80	0.48	0.46	1.00
	In this paper							
C all	6.34	7.79	-0.64	0.11	0.81	0.83	0.60	0.38
C G10	5.17	8.97	-0.69	0.02	0.58	0.69	0.89	0.50

Table C.3: Comparison of FX Style Portfolios: FX Momentum

M refers to our FX momentum portfolio which is based on up to 30 currencies ("all currencies") or the G10 currencies. We show results of momentum strategies that condition on one month past returns ("M 1"), and on the past three ("M 1-3"), six ("M 1-6") and twelve ("M 1-12") months returns. The holding period is one-month. AMP is the FX momentum portfolio as in Asness, Moskowitz, and Pedersen (2012) and is based on the past twelve months returns, skipping the last month return ("2-12"). The AMP portfolio is constructed from the G10 currencies. BER is the FX momentum portfolio as in Burnside, Eichenbaum, and Rebelo (2011) and is based on the past one month returns. The BER portfolio is constructed from up to 20 currencies. The mean is annualized by multiplying by 12, the standard deviation and the Sharpe ratio (SR) are annualized by multiplying by $\sqrt{12}$. The sample period is from 02/1981 to 06/2010.

	FX momentum, sample 02/1981 - 06/2010						
	Mean	Std	Skew	Ac1	SR	Correlation	
AMP M 2-12	2.41	9.52	-0.57	-0.01	0.25	1.00	
BER M 1	4.87	7.16	0.28	0.06	0.68	0.15	1.00
	In this paper						
M 1 all	5.24	7.49	0.27	-0.09	0.70	0.06	0.55
M 1-3 all	5.60	7.70	0.28	-0.08	0.73	0.25	0.43
M 1-6 all	3.74	8.02	0.03	-0.02	0.47	0.47	0.30
M 1-12 all	2.95	8.12	-0.67	-0.00	0.36	0.69	0.17
M 1 G10	2.34	8.53	0.12	-0.02	0.27	0.06	0.53
M 1-3 G10	3.41	8.39	0.15	-0.10	0.41	0.32	0.46
M 1-6 G10	1.45	8.46	-0.48	-0.03	0.17	0.60	0.33
M 1-12 G10	1.20	8.87	-0.28	-0.01	0.14	0.85	0.20

Table C.4: Comparison of FX Style Portfolios: FX Value

"V" refers to our FX value strategy portfolio which is based on up to 30 currencies or the G10 currencies. We show results using the change of the real exchange rate (Q) from $t-2$ to $t-60$ ("2-60"). The table also provides results for 4-60, 7-60 and 13-60 value strategies. In the computation of the value instrument used for ranking currencies, we exclude the past months that overlap with the instrument of a respective momentum strategy. This approach follows Fama and French (1996). Similar to Asness, Moskowitz, and Pedersen (2012), we measure the real exchange rate in $t-60$ using the average of the real exchange rate over the period between $t-54$ and $t-66$. AMP is the FX value portfolio as in Asness, Moskowitz, and Pedersen (2012) and is based on "the 5-year return on the exchange rate, taking into account the interest earned". The AMP portfolio is constructed from the G10 currencies. The mean is annualized by multiplying by 12, the standard deviation and the Sharpe ratio (SR) are annualized by multiplying by $\sqrt{12}$. The sample period is from 02/1981 to 06/2010.

	FX value, sample: 02/1981 - 06/2010					
	Mean	Std	Skew	Ac1	SR	Correlation
AMP V 1-60	4.03	9.02	0.21	0.11	0.45	1.00
	In this paper					
V 2-60 all	4.00	6.74	-0.06	0.14	0.59	0.78
V 4-60 all	4.33	6.76	-0.34	0.10	0.64	0.74
V 7-60 all	4.30	6.49	-0.11	0.04	0.66	0.69
V 13-60 all	4.09	6.53	-0.16	0.02	0.63	0.61
V 2-60 G10	4.83	8.49	0.23	0.11	0.57	0.91
V 4-60 G10	4.99	8.33	-0.04	0.07	0.60	0.86
V 7-60 G10	4.52	8.12	0.13	0.05	0.56	0.80
V 13-60 G10	4.53	8.14	-0.02	0.05	0.56	0.69

Table C.5: Correlation of FX Value and FX Momentum Strategies

Asness et al.	Correlation		In this paper	Correlation	
AMP M 2-12	1	-0.44	M 1-3	1	0.01
AMP V 1-60	-0.44	1	V 4-60	0.01	1

Selecting the baseline FX styles. The construction of the carry strategy is fairly straightforward, but picking a specific FX value and FX momentum strategy is more difficult since there are various ways of computing the instruments (as shown in Tables OA.3 and OA.4). There is no favorite measure in the literature. Menkhoff, Sarno, Schmeling, and Schrimpf (2012b) show that the momentum signal in currency markets is stronger when conditioning on more recent past returns. They find the best results for a MOM(1,1) strategy, which conditions on one-month lagged returns and re-balances the portfolio every month. The authors, however, also report that since past one-month returns are more volatile than longer horizon past returns, transactions costs due to portfolio re-balancing tend to be larger for the strategy conditioning on one-month returns. To take this trade-off into account, we select the momentum strategy conditioning on the past three month returns (FX momentum, M 1-3) as our baseline FX momentum strategy in the main paper.

Furthermore, to avoid correlation between the momentum and value strategies (see Table C.5), it seems sensible to select two strategies with non-overlapping instruments (e.g. as in Fama and French (1996)). Thus, we pick the 5-year change of the real exchange rate skipping the past three months (FX value, V 4-60) as our baseline FX value strategy.

For robustness, we report mean-variance efficiency tests for the alternative momentum and value FX strategies in Section C.4 of this Appendix.

Full sample statistics and transaction costs. An overview of the baseline FX strategies is given in C.6. This Table provides descriptive statistics for the full sample (02/1981-12/2011). Results are shown for strategies based on all currencies and the smaller set of G10 currencies. We also adjust for bid-ask spreads to account for transaction costs.

Table C.6: Descriptive Statistics of FX Investment Styles

Panel A reports results for rank-based FX style portfolios. Panel B reports results using equally weighted long and short positions of currencies. These portfolios are long in the 2/9 quantile with the largest realization of the instrument, and short in the 7/9 quantile. The mean is annualized by multiplying by 12, the standard deviation and the Sharpe ratio (SR) are annualized by multiplying by $\sqrt{12}$. The ba/loss is the percentage loss of the mean return due to our transaction costs adjustment and takes into account portfolio rebalancing using bid ask quotes. The sample period is from 02/1981 to 12/2011.

Panel A: Rank-based portfolio formation (**used in the main paper**)

	Mean	Std	Skew	Ac1	SR	Mean	Std	Skew	Ac1	SR	ba/loss
	All currencies										
	without b-a spreads					with accounting for b-a spreads					
C	6.18	7.50	-0.63	0.10	0.82	5.67	7.52	-0.63	0.10	0.75	-8.29
M 1-3	5.34	7.68	0.25	-0.09	0.70	3.80	7.76	0.24	-0.08	0.49	-28.94
V 4-60	4.18	6.69	-0.31	0.09	0.62	3.82	6.69	-0.32	0.09	0.57	-8.64
	G10 currencies										
	without b-a spreads					with accounting for b-a spreads					
C	5.05	8.76	-0.70	0.03	0.58	4.71	8.78	-0.69	0.03	0.54	-6.88
M 1-3	3.18	8.41	0.13	-0.11	0.38	1.90	8.49	0.12	-0.10	0.22	-40.16
V 4-60	4.51	8.27	-0.00	0.07	0.54	4.19	8.27	-0.00	0.07	0.51	-7.00

Panel B: Equally weighted long short portfolios (2/9-7/9 quantiles)

	Mean	Std	Skew	Ac1	SR	Mean	Std	Skew	Ac1	SR	ba/loss
	All currencies										
	without b-a spreads					with accounting for b-a spreads					
C	7.63	9.28	-0.64	0.10	0.82	6.43	9.28	-0.67	0.10	0.69	-15.78
M 1-3	6.54	9.85	0.51	-0.06	0.66	4.84	9.91	0.50	-0.05	0.49	-25.99
V 4-60	4.96	8.68	-0.35	0.12	0.57	3.91	8.66	-0.38	0.13	0.45	-21.16
	G10 currencies										
	without b-a spreads					with accounting for b-a spreads					
C	5.98	10.92	-0.82	0.03	0.55	5.05	10.93	-0.83	0.03	0.46	-15.51
M 1-3	5.36	10.15	0.09	-0.13	0.53	4.01	10.21	0.09	-0.12	0.39	-25.20
V 4-60	4.62	10.18	-0.09	0.05	0.45	3.85	10.16	-0.09	0.05	0.38	-16.73

Optimized style portfolios. Brandt, Santa-Clara, and Valkanov (2009) propose to directly model the weight in each asset as a function of the specific characteristic.[78] The optimal weights are parameterized as a simple linear function

$$w_{j;t}^{z(s)} = \bar{w}_{j;t} + \frac{1}{J_t}\theta\hat{z}\left(s\right)_{j;t},$$

where $\bar{w}_{j;t}$ is the weight of currency j in an equally weighted benchmark portfolio, $\hat{z}\left(s\right)_{j;t}$ is the cross-sectionally standardized characteristic (zero mean and unit standard deviation), and θ is a coefficient to be estimated.

In the main paper, we rely on the more simple rank-based FX style portfolio which should be more conservative with respect to our results. Nevertheless, it is interesting to know how close our rank-based portfolios are compared to "optimized" portfolios. Table C.7 shows that in terms of Sharpe ratios, the rank-based portfolios are quite close to the optimized counterparts.

It is important to note, however, that the rank-based portfolios use only real-time information, whereas the optimized portfolios are ex-post optimal and require full-sample information. Also portfolio re-balancing is likely to be larger for the optimized portfolios (and thus transaction costs as well).

We also present a strategy which incorporates all three characteristics jointly, similar to Jorda and Taylor (2012). The estimated coefficients θ are 5.98 for the forward premium, 4.20 for momentum, and 3.27 for our value measure, which indicate economically large joint predictability by all three characteristics. Interestingly, combining all three characteristics results in a FX style portfolio with a very large Sharpe ratio and a skewness close to zero. Combining different FX styles seems to able to diversify away the skewness of pure carry trade strategies.

Table C.7: Parametric FX Style Policies

The mean is annualized by multiplying by 12, the standard deviation and the Sharpe ratio (SR) are annualized by multiplying by $\sqrt{12}$. The sample period is from 02/1981 to 12/2011.

	Mean	Std	Skew	Ac1	SR
	without b-a spreads				
Carry trade	20.09	19.71	-0.39	0.19	1.02
FX momentum 1-3	10.56	15.50	1.01	0.01	0.68
FX value 4-60	10.06	13.88	0.28	0.10	0.72
All characteristics	27.86	25.25	0.09	0.05	1.10

[78] See Barroso and Santa-Clara (2012) for a comprehensive application to FX markets.

Options-based Hedging Portfolio for FX Styles

We were able to acquire end of month options data from J.P. Morgan for our set of G10 currencies for the period from 07/1997 to 12/2011. The maturity of each OTC traded option is one month. For each currency and time observation we have five option quotes, namely, 10 delta calls, 25 delta calls, at-the-money options, 25 delta puts, and 10 delta puts. It is a convention in currency option markets that quotes are not reported with a specific strike (price), but instead by their delta (implied volatility) given by the Garman and Kohlhagen (1983) formula. The delta of an option measures the sensitivity of the option price with respect to the spot exchange rate. Briefly, 25 delta options are "out-of-the money", and 10 delta options are "far out-of-the-money" calls or puts. We use implied volatility quotes and the spot exchange rates in our database together with the Garman and Kohlhagen (1983) formula to obtain option prices. See e.g. Jurek (2009), Sarno, Corte, and Tsiakas (2011), and Burnside (2011) for a more detailed discussion of currency option markets.

The construction of the hedged FX style follows Burnside (2011). The payoff for a position in a currency call is

$$RC_{j;t+1} = max\left(0,\, S_{j;t+1} - K_{j;t}\right) - C_{j;t}\left(1 + r_t\right),$$

where $S_{j;t+1}$, $K_{j;t}$, and $C_{j,t}$ denote the currency spot rate, strike, and the price of the call option. For the domestic interest rate, r_t, we use the one month eurodollar rate, available via FRED/St. Louis FED. In the same vein, the payoff for a currency put is

$$RP_{j;t+1} = max\left(0,\, K_{j;t} - S_{j;t+1}\right) - P_{j;t}\left(1 + r_t\right),$$

where $P_{j;t}$ denotes the price of the put option. The payoff for the options-based hedging portfolio is

$$OH_{t+1} = \sum_{j=1}^{J} w_{j;t}^{+} RP_{j;t+1} + \sum_{j=1}^{J} w_{j;t}^{-} RC_{j;t+1},$$

where $w_{j;t}^{+}$ collects the positive positions of a particular FX style and is zero otherwise, and $w_{j;t}^{-}$ collects the negative positions of a particular FX style and is zero otherwise. Finally, the return of the hedged FX style is given by the following expression

$$RZ_{t+1}^{oh} = RZ_{t+1} + OH_{t+1},$$

where RZ_{t+1} is the return of the unhedged FX style. For the subsample where currency options data are available to us, only the carry trade strategy shows a negative skewness and thus we only consider the options-based hedge for this FX style. We do not account for transaction costs for the options-based hedging portfolios. Given our currency options data set, we can build a 10 delta hedging portfolio which hedges against "large" losses only, and a 25 delta hedging portfolio which also hedges against "intermediate" losses.

Overview of Benchmark Assets

Equity. We use the MSCI Total Return (Standard Country) indices for Australia, Canada, Germany, Japan, New Zealand, Norway, Sweden, Switzerland, the United Kingdom, and the United States available via Thomsen Reuters Datastream. These are measured in USD and are available for the period from 01/1981 to 12/2011. The total return index for New Zealand starts in 12/1987. We use observations of the MSCI price index, which starts in 01/1981 already, to measure the New Zealand equity market for the period before 12/1987.

Bonds. There are no official bond market indices covering the time period and the cross-section of countries required for our study. Following Campbell, Sunderam, and Viceira (2010), we therefore use the log yield-return approximation, described by Campbell, Lo, and MacKinlay (1997) in detail, to derive bond market indices. First, long-term government bond yields are taken from the IMF's International Financial Statistics (IFS) database for the same ten countries as for the equity markets. Second, we use the approximated log bond returns to calculate bond indices denominated in local currency. Finally, the local currency bond indices are multiplied with the spot exchange rates for global bond indices denominated in USD.

Excess Returns. Equity and bond market returns are monthly simple returns, based on the indices denominated in USD, in excess of the U.S. Treasury bill rate from Ibbotson which is available on the web site of Kenneth R. French.

Table C.8: Characteristics of International Bonds and International Stocks

The mean is annualized by multiplying by 12, the standard deviation is annualized by multiplying by $\sqrt{12}$. The sample period is from 02/1981 to 12/2011.

	International bonds				International stocks			
	unhedged		fully hedged		unhedged		fully hedged	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std
Australia	6.48	14.30	3.23	7.56	7.94	23.78	4.69	17.31
Canada	6.19	10.76	4.62	7.65	6.28	20.20	4.70	16.39
Germany	4.58	12.60	3.80	5.12	7.87	23.43	7.09	21.62
Japan	4.81	14.76	4.05	6.63	3.47	22.42	2.72	19.61
New Zealand	6.52	16.24	2.27	10.18	4.57	25.40	0.33	20.78
Norway	5.48	12.25	2.98	7.35	9.15	27.31	6.65	24.49
Sweden	4.98	12.91	4.10	6.81	13.47	26.27	12.59	24.42
Switzerland	3.34	13.23	2.66	4.93	8.24	18.02	7.57	16.84
U.K.	4.69	13.07	3.75	6.95	6.49	18.76	5.55	16.32

Value and momentum "everywhere" portfolios. In the robustness section of our paper, we consider a set of popular managed portfolios as benchmark assets, namely value and momentum portfolios for other asset classes. These portfolios cover stock market value and momentum styles for the U.S., U.K., Europe, Japan, and country indices as well as value and momentum styles for commodities and are studied by Asness, Moskowitz, and Pedersen (2012). The data are available on the website of Tobias J. Moskowitz. We collect their "factor" portfolios, which rely on the rank-based weighting to represent a particular strategy. Our FX style portfolios are build on the same rank-based weighting. An important difference to our FX style portfolios is that the Asness, Moskowitz, and Pedersen (2012) value and momentum portfolios are not adjusted for transaction costs due to portfolio re-balancing.

Table C.9: Value and Momentum "Everywhere" Portfolios

This table presents the risk and return characteristics of the momentum and value "everywhere" portfolios as in Asness, Moskowitz, and Pedersen (2012). The mean is annualized by multiplying by 12, the standard deviation and the Sharpe ratio (SR) are annualized by multiplying by $\sqrt{12}$.

	Mean	Std	Skew	Ac1	SR
	Rank-based value "everyhwere" portfolios				
U.S. stocks	3.59	15.29	-0.59	0.10	0.24
U.K. stocks	5.66	14.59	-0.41	0.26	0.39
Europe stocks	5.51	9.69	0.46	0.23	0.57
Japan stocks	10.66	13.39	-0.08	0.21	0.80
Country stock market indices	5.74	9.21	0.18	0.13	0.62
Commodities	4.54	21.25	-0.00	-0.03	0.21
	Rank-based momentum "everyhwere" portfolios				
U.S. stocks	6.28	18.10	-0.04	-0.00	0.35
U.K. stocks	8.71	16.43	-0.28	0.28	0.53
Europe stocks	8.56	13.74	-0.68	0.15	0.62
Japan stocks	1.30	17.64	-0.44	0.10	0.07
Country stock market indices	6.79	11.55	-0.24	-0.02	0.59
Commodities	9.29	20.25	-0.12	0.01	0.46

C.2 Dissecting Currency Hedging and Speculation

This section provides the analytical background for dissection hedging and speculation positions in the mean-variance framework (see also, e.g., Jorion (1994)). The mean-variance solution for an optimal portfolio weight $\mathbf{w}^{\star}$ is

$$\mathbf{w}^{\star} = \lambda \mathbf{w} = \lambda \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}, \tag{43}$$

where $\boldsymbol{\mu}$ is the vector of expected excess returns with covariance matrix $\boldsymbol{\Sigma}$. For the tangency portfolio, we simply have to set λ such that the the elements of $\mathbf{w}^{\star}$ sum to one. The vector of returns contains benchmark assets and test assets, and can be partitioned as

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_R \\ \boldsymbol{\mu}_X \end{pmatrix}, \ \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{RR} & \boldsymbol{\Sigma}_{RX} \\ \boldsymbol{\Sigma}_{XR} & \boldsymbol{\Sigma}_{XX} \end{pmatrix}, \tag{44}$$

where $\boldsymbol{\mu}_R$ denotes the returns of the benchmark assets, and $\boldsymbol{\mu}_X$ denotes the returns of the test assets. In the literature on optimal hedging, the benchmark assets are also called core assets, or underlying assets, and the test assets are called hedging assets. The multiple regression coefficients of the benchmark assets on the test assets are

$$\mathbf{B} = \boldsymbol{\Sigma}_{XX}^{-1} \boldsymbol{\Sigma}_{RX}. \tag{45}$$

The multiple regression coefficients multiplied by minus one are the "optimal" risk management demands of the test assets (e.g. Campbell, Sunderam, and Viceira (2010)). The allocation $\boldsymbol{\mu}_R - \mathbf{B}' \boldsymbol{\mu}_X$ minimizes the risk (variance) of the benchmark assets using the test assets. The covariance matrix of the benchmark assets given the risk management portfolio is

$$\boldsymbol{\Sigma}_{\varepsilon\varepsilon} = \boldsymbol{\Sigma}_{RR} - \mathbf{B}' \boldsymbol{\Sigma}_{XX} \mathbf{B}. \tag{46}$$

The partitioned inverse of the covariance matrix can be used to find an expression for the optimal portfolio weights, which is:

$$\lambda \begin{pmatrix} \mathbf{w}_R \\ \mathbf{w}_X \end{pmatrix} = \lambda \begin{pmatrix} \boldsymbol{\Sigma}_{\varepsilon\varepsilon}^{-1} (\boldsymbol{\mu}_R - \mathbf{B}' \boldsymbol{\mu}_X) \\ \boldsymbol{\Sigma}_{XX}^{-1} \boldsymbol{\mu}_X - \mathbf{B} \mathbf{w}_R \end{pmatrix}. \tag{47}$$

This expression provides a nice interpretation. First, the optimal test asset positions $\mathbf{w}_X$ are determined by a purely speculative component specified by the multivariate Sharpe ratio of test asset returns, $\boldsymbol{\Sigma}_{XX}^{-1} \boldsymbol{\mu}_X$. Second, there is a hedging component given by the risk management demand $-\mathbf{B}$ adjusted by the optimal allocation of benchmark assets $\mathbf{w}_R$.

Thus, when a set of test assets are added to the benchmark assets, there will be diversification benefits due to a hedging and a speculative component. We can see these components better by multiplying the optimal portfolio weights with the returns:

$$\boldsymbol{\mu}' \lambda \mathbf{w} = (\boldsymbol{\mu}_R - \mathbf{B}' \boldsymbol{\mu}_X)' \lambda \mathbf{w}_R + \boldsymbol{\mu}_X' \left(\lambda \boldsymbol{\Sigma}_{XX}^{-1} \boldsymbol{\mu}_X \right),$$

where $-\mathbf{B} \lambda \mathbf{w}_R$ are the hedging positions in the test assets, and $\left(\lambda \boldsymbol{\Sigma}_{XX}^{-1} \boldsymbol{\mu}_X \right)$ are the speculative positions in the test assets.

In the main paper, we test the mean-variance efficiency of global bonds and equities (the benchmark assets) against a portfolio augmented by FX investment styles (the test assets). When these benchmark assets are fully hedged, or optimally hedged using single currency positions, we test for the hedging *and* the speculation benefits of FX investment styles. The results above also verify that if the benchmark assets are hedged against FX style risk ($\boldsymbol{\mu}_R - \mathbf{B}'\boldsymbol{\mu}_X$), i.e. we a priori attribute the hedging component of the test assets to the benchmark, we test exclusively for the speculative benefits of the FX investment styles.

C.3 Stochastic Dominance Tests

It is well known that portfolios based on carry trade strategies exhibit negative skewness and are prone to sudden large losses as documented by Brunnermeier, Nagel, and Pedersen (2009) and Farhi, Fraiberger, Gabaix, Rancière, and Verdelhan (2009). This raises the question of how robust the mean-variance criterion is for optimal portfolio decisions, since this framework does not take higher moments of returns into account.[79]

An appealing framework to avoid the shortcomings of the mean-variance approach is the concept of stochastic dominance. A portfolio is second-order stochastic dominance (SSD) efficient if it is optimal for a non-satiable and risk-averse investor, and it is third-order stochastic dominance (TSD) efficient if it is optimal for a non-satiable, risk-averse and skewness-loving investor (Levy (2006)). Formally, SSD can be represented by any utility function with a non-negative first derivative ($U' \geq 0$) and a non-positive second derivative ($U'' \leq 0$), where the inequalities are strict at least at one point. TSD adds the restriction of a non-negative third derivative of the utility function ($U''' \geqslant 0$) corresponding to the skewness-loving property. Thus, the exact specification of the utility function (and hence the stochastic discount factor) of the investor is left unspecified, but it is merely restricted to be economically sensible.

We follow the testing approach of Post and Versijp (2007). They propose a formal non-parametric multivariate SSD and TSD test. The intuition behind the tests is equivalent to the traditional mean-variance efficiency tests: Once a SDF in line with the SSD / TSD criterion is found, a positive pricing error of a test asset can be interpreted as an investor's desire to increase her allocation with respect to the benchmark. If the pricing error is also statistically significant, the hypothesis of stochastic dominance efficiency of the benchmark allocation is rejected.

Test Statistics. We briefly describe the SSD and TSD tests of Post and Versijp (2007), and refer to their paper for details. The pricing errors of the N test assets are denoted as $\boldsymbol{\alpha}(m) = E(\mathbf{RZ}_\theta m_\theta)$, where m_θ is a candidate SDF, and $\mathbf{RZ}_\theta$ are the test asset excess returns (FX styles). We use subscripts $\theta = 1, ..., \Theta$ to emphasize that the T time-series elements of $\mathbf{RZ}_\theta$ are ranked according to the benchmark portfolio returns which, in turn, are sorted in an increasing order, that is $\mathbf{R}_1'\mathbf{w} < \mathbf{R}_2'\mathbf{w} < ... < \mathbf{R}_\Theta'\mathbf{w}$, where $\mathbf{R}$ are K benchmark asset returns (global bonds and global stocks), and $\mathbf{w}$ are evaluated portfolio weights which generate a stochastic dominance-efficient benchmark portfolio. In our empirical implementation, we minimize the benchmark portfolio's second-order lower partial moment with a target rate of zero, where the lower partial moment is defined as,

$$LPM_2(0) = \frac{1}{T}\sum_{t=1}^{T}\left[max\left(0,(0-R_t)\right)\right]^2, \qquad (48)$$

[79] Note that the stochastic properties of the benchmark assets show pronounced negative skewness as well.

to find valid weights $\mathbf{w}$. Post and Versijp (2007) discuss that minimizing the $LPM_2(0)$ should produce SSD and TSD-efficient portfolios. Similar to the well-known J-test in the GMM framework, the test statistic for SSD efficiency of the benchmark portfolio can be calculated by

$$J_{SSD} = \min_{m \,\epsilon\, M_{SSD}} \Theta \hat{\boldsymbol{\alpha}}(m)' \hat{\boldsymbol{\Omega}}(m)^{-1} \hat{\boldsymbol{\alpha}}(m), \tag{49}$$

where M_{SSD} represents the subset of marginal utility functions that are in line with the SSD criterion, for which the mean of the SDFs (or marginal utility) equals unity, and $\hat{\Omega}(m)$ is the sample covariance matrix of $\hat{\boldsymbol{\alpha}}(m) = \frac{1}{\Theta}\sum_{\theta=1}^{\Theta} \boldsymbol{RZ}_\theta m_\theta$. Given the ordering of the data, M_{SSD} can be represented as the following restrictions on the SDFs

$$M_{SSD} = \left\{ m \,\epsilon\, \mathbb{R}_+^\Theta : \frac{1}{\Theta}\sum_{\theta=1}^{\Theta} m_\theta = 1;\; m_{\theta-1} \geq m_\theta,\; \theta = 2, ..., \Theta \right\}, \tag{50}$$

for the minimization problem in (49), and corresponds to decreasing or at least constant change in marginal utility from low to high returns ($U' \geq 0$ and $U'' \leq 0$). The test statistic for TSD efficiency can be calculated in a similar fashion as

$$J_{TSD} = \min_{m \,\epsilon\, M_{TSD}} \Theta \hat{\boldsymbol{\alpha}}(m)' \hat{\boldsymbol{\Omega}}(m)^{-1} \hat{\boldsymbol{\alpha}}(m), \tag{51}$$

where M_{TSD} represents the subset of marginal utility functions that are in line with the TSD criterion, and for which the mean of the SDFs equals unity. Given the ordering of the data, M_{TSD} can be represented as a set of restrictions on the SDFs for the minimization problem in (51), given by

$$M_{TSD} = \left\{ m \,\epsilon\, M_{SSD} : \frac{m_{\theta-1} - m_{\theta-2}}{\mathbf{R}'_{\theta-1}\mathbf{w} - \mathbf{R}'_{\theta-2}\mathbf{w}} \leq \frac{m_\theta - m_{\theta-1}}{\mathbf{R}'_\theta\mathbf{w} - \mathbf{R}'_{\theta-1}\mathbf{w}},\; \theta = 3, ..., \Theta \right\}, \tag{52}$$

and corresponds to decreasing marginal utility at a diminishing rate from low returns to high returns ($U' \geq 0$, $U'' \leq 0$, and $U''' \geqslant 0$).

Computing J_{SSD} and J_{TSD} is a quadratic minimization problem with linear constraints and can be solved iteratively. We use the initial weighting matrix $\hat{\boldsymbol{\Omega}}(m = 1)$ as proposed by Post and Versijp (2007), and use a two-step estimator as described therein. Post and Versijp (2007) show that the SSD and the TSD test statistics asymptotically follow a central chi-square distribution with N degrees of freedom. Their simulation study of the SSD and TSD test statistics suggests that the asymptotic distribution is appropriate for the sample length used in our analysis.

Table C.10: Stochastic Dominance Efficiency Tests for FX Investment Styles

The table shows pricing errors and p-values for the multivariate second-order stochastic dominance (SSD) test and the multivariate third-order stochastic dominance (TSD) test for efficiency of the mean-$LPM_2(0)$ tangency portfolio of currency risk unhedged/fully hedged global bonds and stocks (covering Australia, Canada, Germany, Japan, New Zealand, Norway, Sweden, Switzerland, U.K., U.S.) relative to FX investment styles (Post and Versijp (2007)). The SSD and TSD test are restricted to price the benchmark portfolio correctly, the pricing kernels and pricing errors are based on one iteration, while the p-values are computed based on the resulting weighting matrix. The test assets are three FX investment styles adjusted for transaction costs and are based on a broad set of up to 30 currencies. The sample period is 02/1981 - 12/2011.

	Benchmark: Global bonds and global stocks							
	Pricing errors (s.e.), unhedged				Pricing errors (s.e.), fully hedged			
	SSD		TSD		SSD		TSD	
Carry trade	0.44	(0.12)	0.52	(0.12)	0.43	(0.12)	0.51	(0.11)
FX momentum	0.30	(0.12)	0.33	(0.12)	0.31	(0.12)	0.31	(0.11)
FX value	0.28	(0.11)	0.33	(0.10)	0.25	(0.11)	0.29	(0.10)
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

C.4 Additional Results

We start the section on additional results by presenting the in-sample properties of the different hedging schemes used in the main paper. After reporting the optimal risk minimizing positions for global bonds and stocks in single currencies and the three FX investment styles, we present tests to assess the economic significance of the hedging schemes. The out-of-sample counterpart of these tests of the hedging schemes is reported in the main paper.

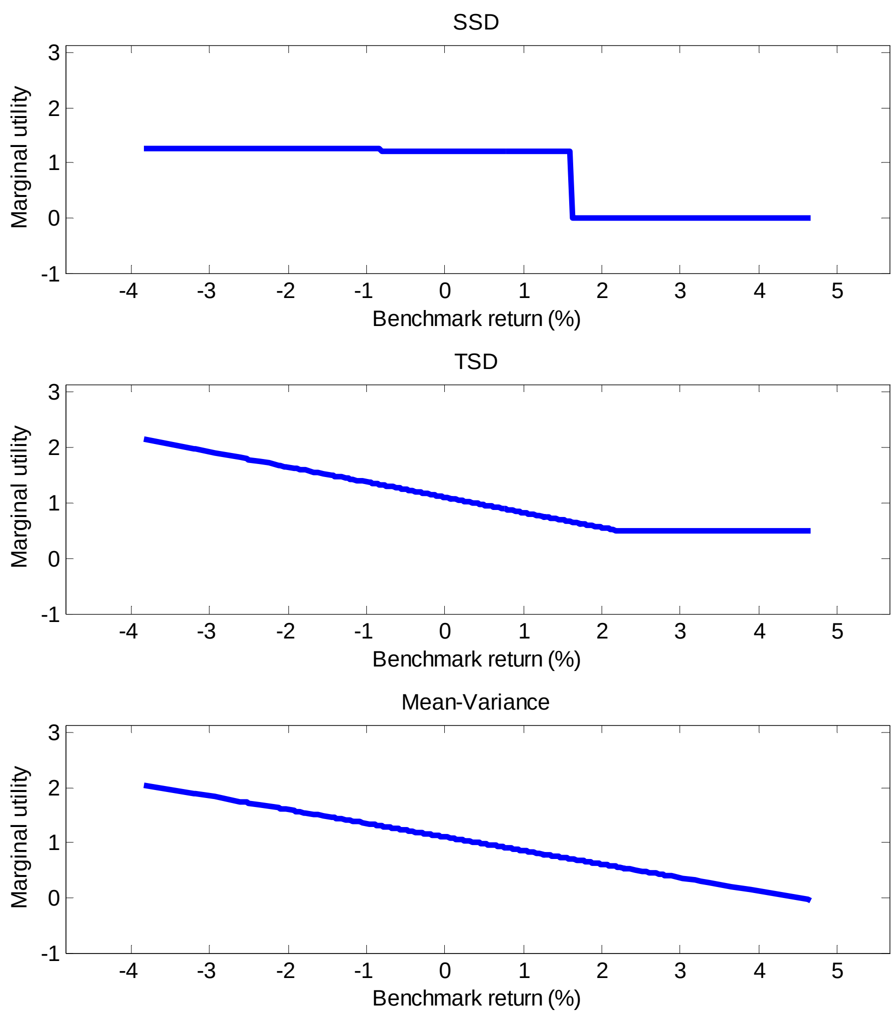
We then turn to additional in-sample tests to assess the diversification benefits of the three FX investment styles. These tables follow the baseline mean-variance efficiency tests of Table 2 in the main text, however, we change various details in the specifications reported here: We report tests relying on bonds and stocks separately as benchmark assets, we test diversification benefits of individual currency positions beyond hedging (as in Glen and Jorion (1993)), we test FX styles based on the G10 currencies without adjusting for transaction-costs, we test two of the three FX investment styles jointly, we compare the effects of further hedging schemes, and we test some alternative versions of the FX momentum and FX value strategies.

Next, we provide additional results for out-of-sample tests. A figure present the estimated out-of-sample portfolio weights of benchmark assets and the FX investment styles which are behind Table 4 of the main paper (our baseline results). The following tables show how results change when we apply different portfolio formation rules (we use the same rules as in DeMiguel, Garlappi, and Uppal (2009)), and the out-of-sample portfolio characteristics when the FX styles are restricted to G10 currencies (to accompany our baseline results in Table 5 of the main paper).

Finally, this section shows additional results for the robustness section of the main paper. We report (in-sample) tests with momentum and value "everywhere" portfolios (Asness, Moskowitz, and Pedersen, 2012) as the benchmark when our FX styles are not adjusted for transaction costs as well (at eye level tests). Afterwards, we present the characteristics for FX styles using multiple characteristics to form rank-based portfolios, and we show the corresponding mean-variance efficiency tests for these alternative FX styles.

Figure C.2: Pricing Kernels

The figure shows the fitted pricing kernels (marginal utility) for the SSD and TSD tests for the sample in Table C.10. We use the $\text{LPM}_2(0)$ tangency portfolio of global bonds and stocks (fully hedged) as benchmark and the FX investment styles carry trade, FX momentum, and FX value as test assets. The pricing kernels are based on one iteration. Below is also the hypothetical pricing kernel of a mean-variance test, implying a quadratic utility function, with the same benchmark and same test assets as for the SSD and TSD tests.



Currency Hedging: In-sample

Table C.11: Risk Management Demands: Global Markets

This table reports estimated risk management demands ($\widetilde{\mathbf{\Psi}}_{RM}$) for GDP PPP weighted global bonds and global stocks as described by Campbell, Sunderam, and Viceira (2010). The hedges include positions in the single G10 currencies and the three FX investment style (carry trade, FX momentum, and FX value; based on all currencies, after transaction costs). The risk management positions minimize the standard deviation of the bond and stock market portfolios (by construction). The sample period is from 02/1981 to 12/2011.

	single currency positions									FX styles		
	AUD	CAD	JPY	EUR	NZD	NOK	SEK	CHF	GBP	C	M	V
	$-\mathbf{B_{RX}}$									$-\mathbf{B_{RZ}}$		
Global bonds (GDP PPP weighted portfolio)												
	optimal hedge											
$\widetilde{\mathbf{\Psi}}_{RM}$	0.03	-0.09	-0.21	-0.11	0.01	0.15	0.06	-0.01	0.02			
t	[0.84]	[-1.68]	[-2.43]	[-3.47]	[0.36]	[2.69]	[1.00]	[-0.12]	[0.33]			
	conditional optimal hedge											
$\widetilde{\mathbf{\Psi}}_{RM}$	0.03	-0.09	-0.21	-0.10	0.00	0.14	0.06	0.00	0.01	0.04	0.02	-0.01
t	[0.77]	[-1.72]	[-2.36]	[-3.20]	[0.10]	[2.50]	[1.05]	[0.04]	[0.20]	[0.99]	[0.55]	[-0.31]
Global stocks (GDP PPP weighted portfolio)												
	optimal hedge											
$\widetilde{\mathbf{\Psi}}_{RM}$	-0.16	-0.63	0.07	-0.06	-0.13	-0.10	-0.30	0.50	0.08			
t	[-1.97]	[-5.64]	[0.41]	[-0.91]	[-1.58]	[-0.60]	[-1.81]	[4.35]	[0.87]			
	conditional optimal hedge											
$\widetilde{\mathbf{\Psi}}_{RM}$	-0.16	-0.62	0.04	-0.10	-0.09	-0.06	-0.30	0.46	0.12	-0.21	-0.04	0.07
t	[-1.86]	[-5.75]	[0.24]	[-1.60]	[-1.07]	[-0.35]	[-1.77]	[3.84]	[1.23]	[-1.89]	[-0.47]	[0.58]

Table C.12: Risk Management Demands: Individual Bonds

This table reports estimated risk management demands ($\widetilde{\boldsymbol{\Psi}}_{RM}$) for global bonds as described by Campbell, Sunderam, and Viceira (2010). The hedges include positions in the single G10 currencies and the three FX investment style (carry trade, FX momentum, and FX value; based on all currencies, after transaction costs). The risk management positions minimize the standard deviation of the bond market portfolios (by construction). The sample period is from 02/1981 to 12/2011. The sample period is from 02/1981 to 12/2011.

	single currency positions									FX styles		
	AUD	CAD	JPY	EUR	NZD	NOK	SEK	CHF	GBP	C	M	V
Risk management demands ($\widetilde{\boldsymbol{\Psi}}_{RM}$) of global bonds												
	$-\mathbf{B_{RX}}$									$-\mathbf{B_{RZ}}$		
U.S.	0.05	-0.15	-0.29	-0.12	0.01	0.21	0.06	-0.02	0.04	0.02	0.02	0.01
AU	-0.07	0.03	0.01	-0.08	-0.01	0.08	0.04	-0.02	-0.03	0.04	0.00	-0.15
CA	0.00	-0.16	-0.08	-0.09	0.06	0.16	-0.04	-0.00	-0.06	0.06	0.00	-0.03
JP	0.01	0.00	-0.16	-0.08	-0.05	0.10	0.06	0.06	0.01	0.06	-0.00	-0.06
GE	-0.02	-0.02	-0.16	-0.08	0.07	-0.03	0.12	-0.03	0.01	0.01	0.01	0.01
NZ	0.06	-0.07	-0.17	-0.09	-0.03	0.09	0.05	0.02	0.08	-0.05	-0.01	-0.11
NO	0.04	-0.04	0.10	-0.11	-0.08	0.10	-0.07	0.07	-0.02	0.05	0.02	-0.10
SE	0.02	0.03	-0.01	-0.06	-0.08	0.15	-0.05	0.04	0.01	0.06	0.03	-0.05
SW	0.01	0.02	-0.03	-0.06	-0.03	0.10	-0.02	-0.02	0.04	0.03	-0.03	-0.00
U.K.	0.04	-0.03	-0.13	-0.08	-0.11	0.11	0.11	0.09	-0.10	0.14	0.07	-0.06
	t-values											
	$t_{-\mathbf{B_{RX}}}$									$t_{-\mathbf{B_{RZ}}}$		
U.S.	1.26	-2.20	-2.29	-2.74	0.36	2.78	0.56	-0.20	0.58	0.45	0.48	0.26
AU	-1.28	0.38	0.11	-1.98	-0.26	1.12	0.69	-0.23	-0.48	0.74	0.02	-2.06
CA	0.02	-1.70	-0.59	-2.04	1.41	1.79	-0.46	-0.08	-0.85	1.10	0.06	-0.61
JP	0.35	0.10	-2.09	-2.73	-1.71	2.22	1.50	1.03	0.18	1.63	-0.15	-1.78
GE	-0.34	-0.28	-2.07	-1.98	1.24	-0.48	2.13	-0.65	0.17	0.30	0.29	0.18
NZ	1.08	-0.94	-1.48	-1.85	-0.57	1.12	0.81	0.26	1.08	-0.84	-0.23	-1.25
NO	0.95	-0.54	0.79	-2.63	-2.18	1.05	-1.07	0.89	-0.46	1.02	0.38	-1.75
SE	0.31	0.60	-0.11	-1.62	-1.52	2.50	-0.95	0.63	0.11	1.22	0.66	-0.74
SW	0.23	0.39	-0.37	-1.88	-1.11	1.81	-0.37	-0.30	1.02	0.77	-1.10	-0.13
U.K.	1.01	-0.47	-1.27	-1.84	-3.11	1.33	1.84	1.29	-1.62	2.34	1.64	-1.09

Table C.13: Risk Management Demands: Individual Stocks

This table reports estimated risk management demands ($\widetilde{\boldsymbol{\Psi}}_{RM}$) for global stocks as described by Campbell, Sunderam, and Viceira (2010). The hedges include positions in the single G10 currencies and the three FX investment styles, carry trade, FX momentum, and FX value (based on all currencies, after transaction costs). The risk management positions minimize by construction the standard deviation of the stock market portfolios. The sample period is from 02/1981 to 12/2011. The sample period is from 02/1981 to 12/2011.

	single currency positions									FX styles		
	AUD	CAD	JPY	EUR	NZD	NOK	SEK	CHF	GBP	C	M	V
	Risk management demands ($\widetilde{\boldsymbol{\Psi}}_{RM}$) of global stocks											
	$-\mathbf{B_{RX}}$									$-\mathbf{B_{RZ}}$		
U.S.	-0.13	-0.70	-0.10	-0.11	-0.07	-0.08	-0.27	0.48	0.16	-0.20	-0.04	0.03
AU	-0.21	-0.38	0.53	-0.14	-0.23	-0.27	-0.22	0.32	-0.09	-0.14	-0.07	-0.07
CA	-0.25	-0.65	0.34	-0.23	-0.05	-0.05	-0.37	0.33	-0.01	-0.20	-0.05	0.11
JP	-0.13	-0.73	-0.18	0.02	-0.14	0.11	-0.43	0.76	0.19	-0.18	-0.15	0.18
GE	-0.22	-0.46	0.28	-0.09	-0.02	0.03	-0.33	0.24	-0.07	-0.35	0.02	0.14
NZ	-0.16	-0.36	0.61	-0.08	-0.05	-0.35	-0.15	0.13	0.01	-0.17	0.06	-0.08
NO	-0.46	-0.68	0.58	-0.03	-0.15	-0.31	-0.29	0.49	0.00	-0.16	0.08	0.05
SE	-0.38	-0.67	0.28	-0.07	-0.12	-0.23	0.14	0.48	0.04	-0.06	0.03	-0.05
SW	-0.12	-0.31	0.13	-0.10	-0.15	0.09	-0.45	0.62	0.07	-0.14	-0.03	-0.10
U.K.	-0.15	-0.48	0.26	-0.16	-0.19	-0.29	-0.27	0.43	0.34	-0.09	-0.04	0.04
	t-values											
	$t_{-\mathbf{B_{RX}}}$									$t_{-\mathbf{B_{RZ}}}$		
U.S.	-1.37	-5.56	-0.46	-1.43	-0.73	-0.45	-1.32	3.70	1.56	-1.51	-0.34	0.23
AU	-2.04	-2.73	2.17	-1.77	-1.79	-1.43	-1.86	2.08	-0.71	-1.10	-0.50	-0.44
CA	-2.61	-4.61	1.70	-2.73	-0.63	-0.31	-1.65	2.43	-0.05	-1.44	-0.55	0.83
JP	-0.82	-4.26	-0.68	0.23	-0.95	0.44	-2.69	3.48	1.34	-1.14	-1.11	1.06
GE	-1.51	-2.76	1.12	-0.86	-0.19	0.14	-1.77	1.39	-0.48	-2.27	0.13	0.77
NZ	-1.29	-2.20	2.53	-0.72	-0.43	-1.57	-0.61	0.80	0.06	-1.06	0.49	-0.43
NO	-3.04	-3.44	2.20	-0.24	-1.10	-1.29	-1.65	2.29	0.00	-0.78	0.51	0.29
SE	-2.58	-3.54	0.90	-0.59	-0.97	-0.77	0.41	2.13	0.24	-0.33	0.18	-0.23
SW	-1.13	-2.41	0.59	-1.00	-1.27	0.52	-2.52	4.72	0.61	-0.93	-0.26	-0.72
U.K.	-1.53	-3.80	1.20	-2.09	-1.89	-1.56	-1.66	3.25	2.69	-0.68	-0.36	0.25

Table C.14: Economic Significance of Risk Management Demands: In-sample

The table provides standard deviations of global bonds and global stocks using different hedging schemes. The optimal hedge minimizes the standard deviation of the returns using all G10 currencies. The conditional optimal hedge includes positions in the single G10 currencies and the FX investment styles. The sample period is from 02/1981 to 12/2011.

	Standard deviation % $\times\sqrt{12}$				Tests of significance					
	no hedge	full hedge	opti. hedge	cond. opti.	FH vs. NO		OH vs. FH		OH vs. CO	
	(NO)	(FH)	(OH)	(CO)	F-stat	p-val.	F-stat	p-val.	F-stat	p-val.
GDP PPP weighted portfolios										
Global bonds	7.59	5.38	5.04	5.03	1.99	0.00	1.14	0.11	1.00	0.49
Global stocks	15.12	14.31	12.07	11.99	1.12	0.15	1.41	0.00	1.01	0.45
Individual global bond markets										
U.S.		7.31	6.87	6.86			1.13	0.12	1.00	0.49
AU	14.30	7.56	7.43	7.37	3.58	0.00	1.03	0.37	1.02	0.44
CA	10.76	7.65	7.43	7.42	1.98	0.00	1.06	0.28	1.00	0.49
JP	12.60	5.12	4.92	4.89	6.07	0.00	1.08	0.23	1.01	0.45
GE	14.76	6.63	6.30	6.30	4.95	0.00	1.11	0.16	1.00	0.50
NZ	16.24	10.18	10.08	10.05	2.55	0.00	1.02	0.43	1.01	0.48
NO	12.25	7.35	7.12	7.09	2.78	0.00	1.06	0.27	1.01	0.46
SE	12.91	6.81	6.64	6.62	3.60	0.00	1.05	0.32	1.01	0.48
SW	13.23	4.93	4.81	4.80	7.20	0.00	1.05	0.32	1.00	0.49
U.K.	13.07	6.95	6.71	6.62	3.53	0.00	1.07	0.24	1.03	0.40
Individual global stock markets										
U.S.		15.53	13.45	13.39			1.33	0.00	1.01	0.47
AU	23.78	17.31	15.33	15.29	1.89	0.00	1.28	0.01	1.01	0.48
CA	20.20	16.39	14.08	14.00	1.52	0.00	1.36	0.00	1.01	0.46
JP	23.43	21.62	19.46	19.37	1.18	0.06	1.23	0.02	1.01	0.46
GE	22.42	19.61	18.26	18.13	1.31	0.01	1.15	0.09	1.02	0.44
NZ	25.40	20.78	19.84	19.80	1.50	0.00	1.10	0.19	1.00	0.49
NO	27.31	24.49	21.44	21.41	1.24	0.02	1.30	0.01	1.00	0.49
SE	26.27	24.42	22.38	22.37	1.16	0.08	1.19	0.05	1.00	0.50
SW	18.02	16.84	15.06	15.02	1.15	0.10	1.25	0.02	1.01	0.48
U.K.	18.76	16.32	14.40	14.38	1.32	0.00	1.28	0.01	1.00	0.49

Currency Speculation: In-sample

Table C.15: Mean-Variance Efficiency Tests for FX Styles: Benchmark with only Bonds

The table displays mean-variance efficiency tests as in the main text, except that the benchmark assets are only global bonds (K=10). The sample period is 02/1981 - 12/2011.

Benchmark: Global bonds										
	full hedge $\widetilde{\boldsymbol{\Psi}}_{RM} = -\mathbf{1_j}$			optimal hedge $\widetilde{\boldsymbol{\Psi}}_{RM} = -\mathbf{B_{RX}}$			conditional optimal $\widetilde{\boldsymbol{\Psi}}_{RM} = [-\mathbf{B_{RX}}, -\mathbf{B_{RZ}}]'$			
	W	SDF	Sharpe	W	SDF	Sharpe	W	SDF	Sharpe	
	p-value		ratio	p-value		ratio	p-value		ratio	
Panel A: FX styles based on all currencies										
		Bench.:	0.92		Bench.:	1.00			Bench.	+FXS
Carry trade	0.000	0.000	1.32	0.000	0.001	1.28	0.000	0.002	1.05	1.29
FX momentum	0.003	0.005	1.04	0.002	0.004	1.12	0.003	0.005	1.01	1.12
FX value	0.002	0.005	1.07	0.004	0.008	1.14	0.003	0.007	0.99	1.14
All	0.000	0.000	1.49	0.000	0.000	1.46	0.000	0.000	1.04	1.46
Panel B: FX styles based on G10 currencies										
		Bench.:	0.92		Bench.:	1.00			Bench.	+FXS
Carry trade	0.000	0.001	1.18	0.003	0.013	1.14	0.003	0.012	1.05	1.18
FX momentum	0.212	0.215	0.95	0.194	0.201	1.02	0.186	0.187	0.92	0.95
FX value	0.010	0.012	1.04	0.014	0.016	1.11	0.007	0.008	0.91	1.04
All	0.000	0.003	1.27	0.003	0.031	1.23	0.001	0.026	1.04	1.27

Table C.16: Mean-Variance Efficiency Tests for FX Styles:
Benchmark with only Stocks

The table displays mean-variance efficiency tests as in the main text, except that the benchmark assets are only global stocks (K=10). The sample period is 02/1981 - 12/2011.

Benchmark: Global bonds										
	full hedge $\widetilde{\mathbf{\Psi}}_{RM} = -\mathbf{1_j}$			optimal hedge $\widetilde{\mathbf{\Psi}}_{RM} = -\mathbf{B_{RX}}$			conditional optimal $\widetilde{\mathbf{\Psi}}_{RM} = [-\mathbf{B_{RX}}, -\mathbf{B_{RZ}}]'$			
	W	SDF	Sharpe	W	SDF	Sharpe	W	SDF	Sharpe	
	p-value		ratio	p-value		ratio	p-value		ratio	
Panel A: FX styles based on all currencies										
		Bench.:	0.68		Bench.:	0.66			Bench.	+FXS
Carry trade	0.000	0.002	0.97	0.000	0.001	1.01	0.000	0.001	0.68	1.01
FX momentum	0.004	0.005	0.83	0.003	0.004	0.81	0.003	0.003	0.65	0.81
FX value	0.013	0.021	0.84	0.007	0.014	0.84	0.003	0.008	0.62	0.84
All	0.000	0.000	1.16	0.000	0.000	1.21	0.000	0.000	0.62	1.20
Panel B: FX styles based on G10 currencies										
		Bench.:	0.68		Bench.:	0.66			Bench.	+FXS
Carry trade	0.015	0.033	0.82	0.003	0.011	0.85	0.004	0.012	0.69	0.87
FX momentum	0.189	0.194	0.71	0.196	0.197	0.69	0.188	0.193	0.66	0.69
FX value	0.022	0.024	0.82	0.014	0.018	0.81	0.009	0.012	0.63	0.81
All	0.021	0.051	0.93	0.005	0.031	0.97	0.003	0.031	0.66	0.98

Table C.17: Individual Currency Returns

The table displays mean-variance efficiency tests with ten global bonds and ten global stocks as the benchmark assets as in the main paper. In Panel A, we test individual currency returns (G10) as test assets. In Panel B, we test the three FX investment styles (all currencies, adjusted for transaction costs) as test assets and we include the individual currency returns (G10) as benchmark assets. The sample period is 02/1981 - 12/2011.

	W	SDF	Sharpe	W	SDF	Sharpe	
	p-value		ratio	p-value		ratio	
Panel A: Individual currency returns as test assets (K=20)							
	full hedge $\widetilde{\boldsymbol{\Psi}}_{RM} = -\mathbf{1_j}$			optimal hedge $\widetilde{\boldsymbol{\Psi}}_{RM} = -\mathbf{B_{RX}}$			
			1.17			1.20	
RX (N=9)	0.370	0.503	1.30	0.638	0.724	1.30	
Panel B: Individual currency returns as benchmark assets (K=29)							
	optimal hedge $\widetilde{\boldsymbol{\Psi}}_{RM} = -\mathbf{B_{RX}}$			conditional optimal $\widetilde{\boldsymbol{\Psi}}_{RM} = [-\mathbf{B_{RX}}, -\mathbf{B_{RZ}}]'$			
			1.30			K	K+N
Carry trade (N=1)	0.000	0.000	1.56	0.000	0.001	1.36	1.56
FX momentum (N=1)	0.002	0.002	1.40	0.004	0.004	1.31	1.40
FX value (N=1)	0.004	0.006	1.43	0.002	0.003	1.27	1.43
All (N=3)	0.000	0.000	1.72	0.000	0.000	1.34	1.72

Table C.18: FX Styles based on G10 Currencies without Transaction Costs Adjustment

The table displays mean-variance efficiency tests as in the main text. The test assets are FX investment styles based on G10 currencies without accounting for transaction costs. The benchmark assets are global bonds and global stocks. The sample period is 02/1981 - 12/2011.

Benchmark: Global bonds and global stocks										
	full hedge $\widetilde{\boldsymbol{\Psi}}_{RM} = -\mathbf{1_j}$			optimal hedge $\widetilde{\boldsymbol{\Psi}}_{RM} = -\mathbf{B_{RX}}$			conditional optimal $\widetilde{\boldsymbol{\Psi}}_{RM} = [-\mathbf{B_{RX}}, -\mathbf{B_{RZ}}]'$			
	W	SDF	Sharpe	W	SDF	Sharpe	W	SDF	Sharpe	
	p-value		ratio	p-value		ratio	p-value		ratio	
Panel A: FX styles based on G10 currencies (not adjusted for transaction costs)										
		Bench.:	1.17		Bench.:	1.20			Bench.:	+FXS
Carry trade	0.000	0.001	1.37	0.001	0.007	1.34	0.002	0.010	1.25	1.38
FX momentum	0.028	0.031	1.22	0.028	0.026	1.26	0.026	0.025	1.20	1.26
FX value	0.010	0.012	1.27	0.012	0.016	1.30	0.007	0.010	1.18	1.30
All	0.000	0.003	1.46	0.001	0.011	1.44	0.001	0.011	1.21	1.46

Table C.19: Mean-Variance Efficiency Tests for Two FX Styles

The table displays mean-variance efficiency tests as in the main text. The test assets are two FX investment styles based on all currencies (N=2). For comparisons, the table also reports results for joint tests of all three FX investment styles (N=3) as provided in the main paper. The benchmark assets are global bonds and global stocks. The sample period is 02/1981 - 12/2011.

Benchmark: Global bonds and global stocks										
	full hedge $\widetilde{\mathbf{\Psi}}_{RM} = -\mathbf{1_j}$			optimal hedge $\widetilde{\mathbf{\Psi}}_{RM} = -\mathbf{B_{RX}}$			conditional optimal $\widetilde{\mathbf{\Psi}}_{RM} = [-\mathbf{B_{RX}}, -\mathbf{B_{RZ}}]'$			
	W	SDF	Sharpe	W	SDF	Sharpe	W	SDF	Sharpe	
	p-value		ratio	p-value		ratio	p-value		ratio	
Panel A: FX styles based on all currencies										
		Bench.:	1.17		Bench.:	1.20			Bench.	+FXS
Carry + FX mo.	0.000	0.000	1.56	0.000	0.000	1.56	0.000	0.000	1.29	1.57
Carry + FX va.	0.000	0.000	1.54	0.000	0.003	1.53	0.000	0.004	1.23	1.54
FX mo. + FX va.	0.001	0.001	1.35	0.001	0.001	1.39	0.001	0.001	1.18	1.38
All	0.000	0.000	1.61	0.000	0.000	1.61	0.000	0.000	1.25	1.62
Panel B: FX styles based on G10 currencies										
		Bench.:	1.17		Bench.:	1.20			Bench.	+FXS
Carry + FX mo.	0.002	0.007	1.35	0.009	0.025	1.34	0.014	0.036	1.24	1.36
Carry + FX va.	0.000	0.004	1.41	0.001	0.019	1.39	0.001	0.019	1.21	1.41
FX mo. + FX va.	0.053	0.049	1.27	0.057	0.055	1.30	0.039	0.037	1.18	1.30
All	0.001	0.009	1.42	0.004	0.034	1.40	0.003	0.036	1.21	1.41

Table C.20: Other Hedges

The table displays mean-variance efficiency tests as in the main text. The conditional hedge in this table starts with fully hedged benchmark assets and adds risk minimizing positions in the FX investment styles. There are no risk minimizing positions in G10 currencies. The benchmark assets are global bonds and global stocks. The sample period is 02/1981 - 12/2011.

Benchmark: Global bonds and global stocks							
	no hedge $\widetilde{\mathbf{\Psi}}_{RM} = \mathbf{0}$			"conditional" hedge $\widetilde{\mathbf{\Psi}}_{RM} = [-\mathbf{1}_{\mathbf{j}}, -\mathbf{B}_{RZ}]$			
	W	SDF	Sharpe	W	SDF	Sharpe	
	p-value		ratio	p-value		ratio	
Panel A: FX styles based on all currencies							
		Bench.:	1.03			Bench.	+FXS
Carry trade	0.000	0.000	1.40	0.000	0.002	1.26	1.47
FX momentum	0.001	0.002	1.17	0.003	0.004	1.17	1.27
FX value	0.004	0.005	1.18	0.003	0.007	1.13	1.27
All	0.000	0.000	1.59	0.000	0.000	1.25	1.61
Panel B: FX styles based on G10 currencies							
		Bench.:	1.03			Bench.	+FXS
Carry trade	0.000	0.001	1.26	0.003	0.013	1.23	1.34
FX momentum	0.111	0.119	1.07	0.202	0.194	1.17	1.19
FX value	0.010	0.009	1.16	0.012	0.014	1.15	1.26
All	0.000	0.001	1.39	0.002	0.035	1.21	1.42

Table C.21: Mean-Variance Efficiency Tests: Alternative Momentum Strategies

The table displays mean-variance efficiency tests as in the main text. The FX momentum strategies are build on the last month return, as well as the past three, six, and twelve months returns. Descriptive statistics are provided in Table C.3. FX styles are not adjusted for transaction costs. The sample period is 02/1981 - 12/2011.

Benchmark: Global bonds and global stocks										
	full hedge $\widetilde{\mathbf{\Psi}}_{RM} = -\mathbf{1_j}$			optimal hedge $\widetilde{\mathbf{\Psi}}_{RM} = -\mathbf{B_{RX}}$			conditional optimal $\widetilde{\mathbf{\Psi}}_{RM} = [-\mathbf{B_{RX}}, -\mathbf{B_{RZ}}]'$			
	W p-value	SDF p-value	Sharpe ratio	W p-value	SDF p-value	Sharpe ratio	W p-value	SDF p-value	Sharpe ratio	
Panel A: FX styles based on all currencies										
		Bench.:	1.17		Bench.:	1.20			Bench.	+FXS
FX momentum 1	0.001	0.001	1.32	0.000	0.001	1.35	0.000	0.000	1.19	1.35
FX momentum 3	0.000	0.000	1.36	0.000	0.000	1.40	0.000	0.000	1.22	1.40
FX momentum 6	0.013	0.014	1.24	0.013	0.012	1.28	0.016	0.013	1.20	1.28
FX momentum 12	0.032	0.047	1.23	0.044	0.061	1.26	0.123	0.143	1.22	1.25
Panel B: FX styles based on G10 currencies										
		Bench.:	1.17		Bench.:	1.20			Bench.	+FXS
FX momentum 1	0.373	0.393	1.18	0.374	0.391	1.21	0.211	0.219	1.19	1.21
FX momentum 3	0.028	0.031	1.22	0.028	0.026	1.26	0.026	0.025	1.20	1.26
FX momentum 6	0.341	0.325	1.18	0.324	0.304	1.21	0.406	0.378	1.20	1.21
FX momentum 12	0.408	0.408	1.18	0.461	0.452	1.21	0.588	0.567	1.20	1.21

Table C.22: Mean-Variance Efficiency Tests: Alternative Value Strategies

The table displays mean-variance efficiency tests as in the main text. The FX value strategies are build on the change of the real exchange rate from $t-2$ to $t-60$, and 4-60, 7-60, and 13-60. Descriptive statistics are provided in Table C.4. FX styles are not adjusted for transaction costs. The sample period is 02/1981 - 12/2011.

Benchmark: Global bonds and global stocks									
	full hedge $\widetilde{\mathbf{\Psi}}_{RM} = -\mathbf{1_j}$			optimal hedge $\widetilde{\mathbf{\Psi}}_{RM} = -\mathbf{B_{RX}}$			conditional optimal $\widetilde{\mathbf{\Psi}}_{RM} = [-\mathbf{B_{RX}}, -\mathbf{B_{RZ}}]'$			
	W p-value	SDF p-value	Sharpe ratio	W p-value	SDF p-value	Sharpe ratio	W p-value	SDF p-value	Sharpe ratio	
Panel A: FX styles based on all currencies										
		Bench.:	1.17		Bench.:	1.20			Bench.	+FXS
FX value 2-60	0.010	0.013	1.27	0.012	0.016	1.31	0.005	0.008	1.17	1.30
FX value 4-60	0.004	0.007	1.29	0.004	0.009	1.33	0.001	0.004	1.17	1.32
FX value 7-60	0.001	0.003	1.31	0.001	0.003	1.35	0.001	0.001	1.18	1.34
FX value 13-60	0.004	0.006	1.28	0.005	0.007	1.32	0.002	0.004	1.17	1.31
Panel B: FX styles based on G10 currencies										
		Bench.:	1.17		Bench.:	1.20			Bench.	+FXS
FX value 2-60	0.008	0.017	1.27	0.018	0.032	1.28	0.018	0.040	1.21	1.28
FX value 4-60	0.000	0.000	1.41	0.000	0.001	1.40	0.000	0.001	1.19	1.40
FX value 7-60	0.046	0.057	1.23	0.055	0.062	1.26	0.037	0.043	1.19	1.25
FX value 13-60	0.001	0.002	1.34	0.001	0.004	1.35	0.001	0.004	1.19	1.35

Currency Speculation: Out-of-sample

Figure C.3: Time-Variation of Out-of-Sample Portfolio Weights: Full Hedge

The figure shows the mean-variance out-of-sample portfolio weights when the benchmark assets are fully hedged (without short sales). Short-sales are not allowed. This is the time-series of portfolio weights behind the results in the main text of Table 7, Panel B, full hedge, with all FX investment styles.

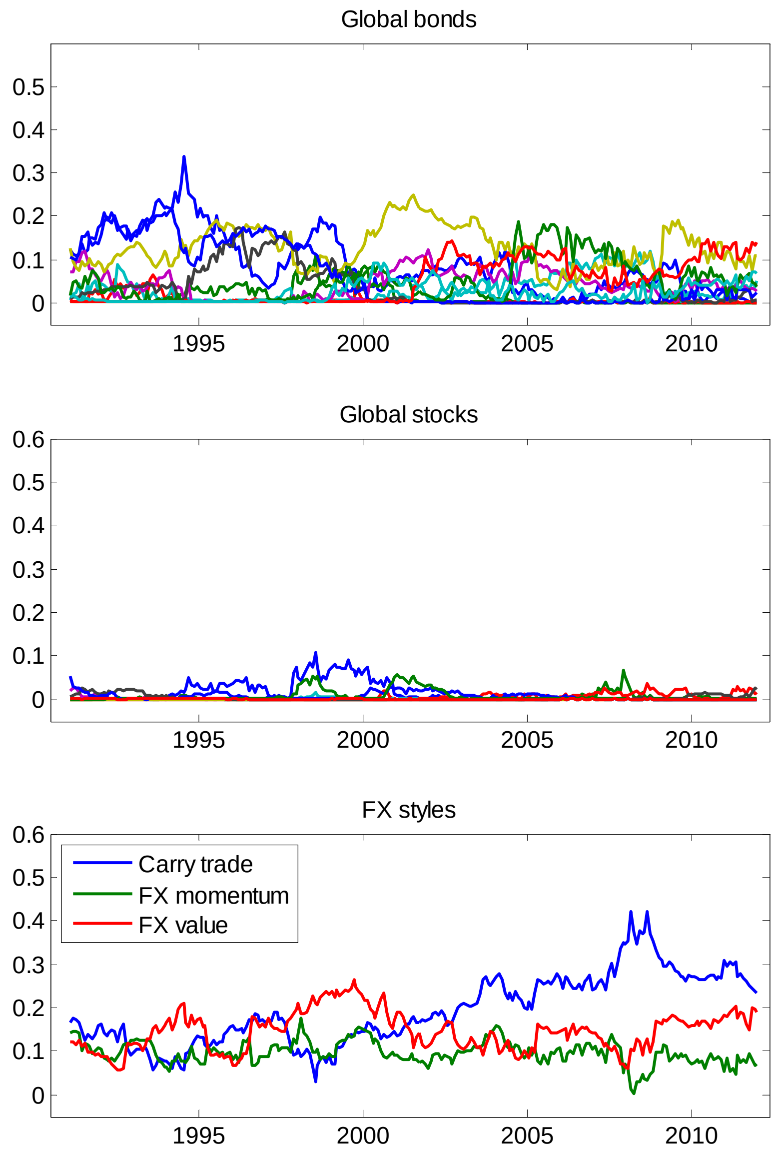


Figure C.4: Time-Variation of Out-of-Sample Portfolio Weights: Full Hedge, with Short Sales

The figure shows the mean-variance out-of-sample portfolio weights when the benchmark assets are fully hedged. In contrast to the specifications in the main text (Table 7, Panel B, full hedge) there are no constraints on the portfolio weights.

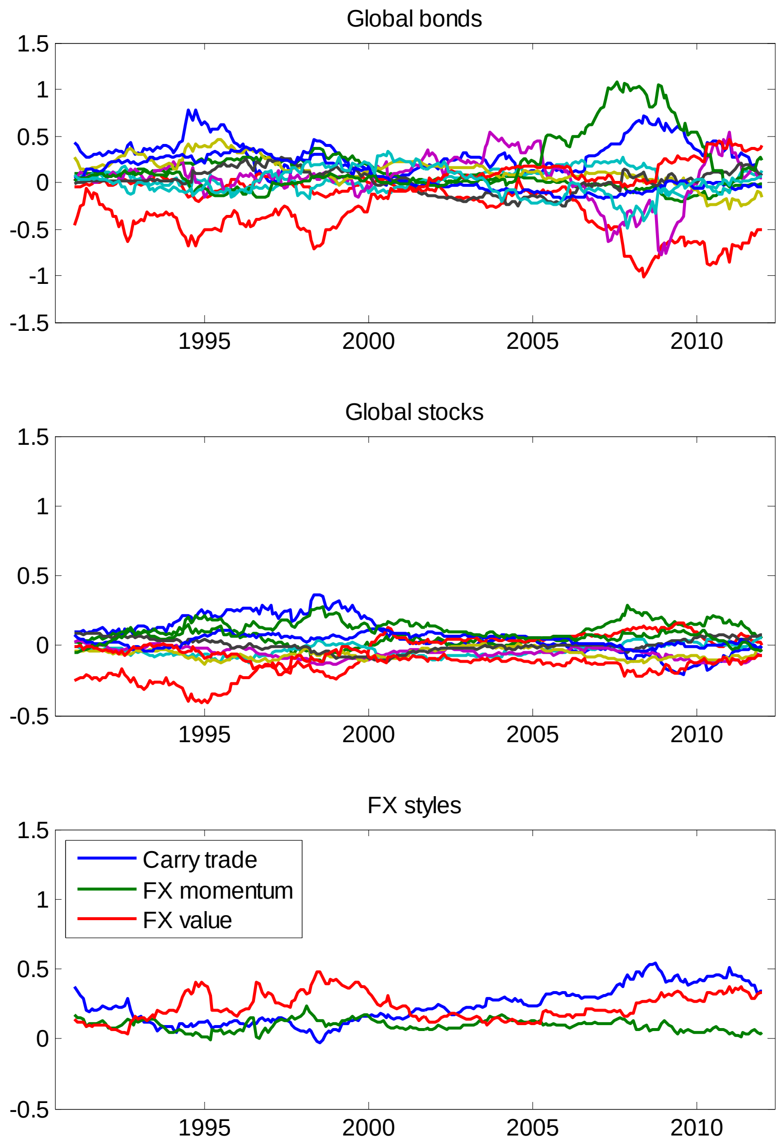


Table C.23: Out-of-Sample Sharpe Ratios: Further Portfolio Formation Rules

The table reports out-of-sample Sharpe ratios as in the main text, except that we apply additional portfolio formation rules which are also used by DeMiguel, Garlappi, and Uppal (2009). The sample period is 02/1981 - 12/2011, all portfolio performance results are based on the sample period 02/1991 - 12/2011.

	full hedge			
	All currencies		G10 currencies	
	Sharpe ratio	$t_{\triangle SR}$	Sharpe ratio	$t_{\triangle SR}$
Panel A:	**1/N** all individual assets			
Benchmark Sharpe ratio:	0.66		0.66	
Carry trade	0.69	[2.63]	0.68	[1.70]
FX momentum	0.69	[2.69]	0.67	[1.18]
FX value	0.69	[2.61]	0.69	[2.01]
All	0.75	[4.43]	0.72	[2.66]
Panel B:	Mean-variance **with** short-sales			
Benchmark Sharpe ratio:	0.56		0.56	
Carry trade	0.83	[3.04]	0.62	[0.87]
FX momentum	0.66	[1.55]	0.55	[-0.05]
FX value	0.68	[1.59]	0.64	[0.94]
All	0.99	[3.91]	0.69	[1.37]
Panel C:	Mean-variance **Bayes-Stein**			
Benchmark Sharpe ratio:	0.72		0.72	
Carry trade	1.04	[2.86]	0.83	[1.43]
FX momentum	0.85	[1.74]	0.73	[0.41]
FX value	0.87	[1.73]	0.82	[1.08]
All	1.25	[4.19]	0.93	[1.92]
Panel D:	**Minimum-variance**			
Benchmark Sharpe ratio:	0.85		0.85	
Carry trade	1.14	[2.55]	0.99	[1.55]
FX momentum	0.99	[1.96]	0.88	[0.47]
FX value	0.96	[1.32]	0.90	[0.74]
All	1.31	[3.57]	1.04	[1.73]

Table C.24: Out-of-Sample Portfolio Characteristics and Positions: G10 Currencies

The table reports portfolio characteristics and average positions in benchmark assets and the three FX investment styles ("+FXS") of the out-of-sample allocations in Table 2 of the main paper. VaR is the Value at Risk. ES is the Expected shortfall and is based on the historical return distribution. Returns are annualized by multiplying with 12, the standard deviation is annualized by multiplying with $\sqrt{12}$.

Panel A: Characteristics of out-of-sample portfolios; FX styles are based on G10 currencies								
	Naive portfolios				Mean-variance without short sales			
	full hedge	optimal hedge	co.opt. hedge	co.opt. + FXS	full hedge	optimal hedge	co.opt. hedge	co.opt. + FXS
Mean	4.49	4.25	2.84	3.15	3.50	3.50	2.60	2.91
Std	7.17	5.25	6.63	4.72	4.59	3.46	4.97	3.63
SR	0.63	0.81	0.43	0.67	0.76	1.01	0.52	0.80
Skew	-0.66	-0.35	0.40	0.48	-0.38	-0.14	-0.46	-0.38
VaR 5%	39.11	27.96	32.52	22.45	20.51	17.24	25.94	16.35
ES 5%	54.64	36.75	43.80	28.64	32.93	23.65	35.97	25.76
Panel B: Out-of-sample portfolio positions								
Global bonds (Σ)	0.50	0.33	0.50	0.33	0.86	0.61	0.81	0.55
Global stocks (Σ)	0.50	0.33	0.50	0.33	0.14	0.07	0.19	0.12
Carry trade	0.00	0.11	0.00	0.11	0.00	0.15	0.00	0.14
FX momentum	0.00	0.11	0.00	0.11	0.00	0.02	0.00	0.02
FX value	0.00	0.11	0.00	0.11	0.00	0.15	0.00	0.17

Table C.25: FX Styles against Value and Momentum "Everywhere":
FX Styles without Transaction Costs Adjustment

The table displays mean-variance efficiency tests as in the main text. The benchmark assets are the Asness, Moskowitz, and Pedersen (2012) value and momentum portfolios "everywhere" (the stock markets of the U.S., the U.K., Europe, Japan, country indices, and commodities). The FX investment styles are based on all currencies in Panel A and they are based on the G10 currencies in Panel B. The test assets and the benchmark assets do not account for transaction costs to provide a level playing field for comparison to the Asness, Moskowitz, and Pedersen (2012) benchmark assets. The sample period is 07/1981 - 06/2010.

Benchmark: Value and momentum "everywhere" (not adjusted for transaction costs)							
	optimal hedge $\widetilde{\mathbf{\Psi}}_{RM} = -\mathbf{B_{RX}}$			conditional optimal $\widetilde{\mathbf{\Psi}}_{RM} = [-\mathbf{B_{RX}}, -\mathbf{B_Z}]'$			
	W	SDF	Sharpe	W	SDF	Sharpe	
	p-value		ratio	p-value		ratio	
Panel A: FX styles based on all currencies (not adjusted for transaction costs)							
		Bench.:	2.01			Bench.	+FXS
Carry trade	0.000	0.001	2.22	0.000	0.004	2.07	2.23
FX momentum	0.000	0.002	2.13	0.000	0.002	2.01	2.13
FX value	0.030	0.056	2.08	0.011	0.027	1.99	2.08
All	0.000	0.000	2.37	0.000	0.001	2.04	2.37
Panel B: FX styles based on G10 currencies (not adjusted for transaction costs)							
		Bench.:	2.01			Bench.	+FXS
Carry trade	0.044	0.079	2.07	0.018	0.045	1.98	2.05
FX momentum	0.127	0.098	2.04	0.077	0.059	2.01	2.04
FX value	0.026	0.044	2.08	0.026	0.035	2.01	2.08
All	0.040	0.133	2.14	0.021	0.081	1.97	2.12

Table C.26: Multi-ranked FX Styles: Characteristics

The table displays characteristics of rank-based FX styles which are build on multiple signals. The signals are the forward discount (C), the last 3-month return (M), and 5-year change of the real exchange rate (V). The mean is annualized by multiplying by 12, the standard deviation and the Sharpe ratio (SR) are annualized by multiplying by $\sqrt{12}$. The sample period is from 02/1981 to 12/2011.

	Mean	Std	Skew	Ac1	SR	Mean	Std	Skew	Ac1	SR	ba/loss
	All currencies										
	without b-a spreads					with accounting for b-a spreads					
C+M	7.14	7.01	-0.16	0.03	1.02	5.94	7.07	-0.16	0.04	0.84	-16.83
C+V	7.11	6.95	-0.57	0.07	1.02	6.54	6.96	-0.57	0.07	0.94	-7.93
M+V	5.86	7.89	0.06	-0.00	0.74	4.56	7.94	0.05	0.01	0.57	-22.28
C+M+V	7.20	6.79	-0.30	0.01	1.06	6.25	6.83	-0.31	0.01	0.92	-13.16
	G10 currencies										
	without b-a spreads					with accounting for b-a spreads					
C+M	4.53	8.22	-0.79	-0.04	0.55	3.56	8.26	-0.79	-0.03	0.43	-21.32
C+V	6.86	8.52	-0.46	0.02	0.81	6.34	8.54	-0.46	0.02	0.74	-7.57
M+V	4.53	8.43	0.39	0.02	0.54	3.44	8.49	0.37	0.03	0.41	-23.99
C+M+V	6.19	8.39	-0.30	0.03	0.74	5.32	8.42	-0.29	0.03	0.63	-14.09

Table C.27: Multi-ranked FX Styles: Mean-Variance Efficiency Tests

The table displays mean-variance efficiency tests as in the main text. The FX styles are build on multiple signals, in particular, the forward discount (C), the last 3-month return (M), and 5-year change of the real exchange rate (V). The sample period is 02/1981 - 12/2011.

Benchmark: Global bonds and global stocks										
	full hedge $\widetilde{\boldsymbol{\Psi}}_{RM} = -\mathbf{1_j}$			optimal hedge $\widetilde{\boldsymbol{\Psi}}_{RM} = -\mathbf{B_{RX}}$			conditional optimal $\widetilde{\boldsymbol{\Psi}}_{RM} = [-\mathbf{B_{RX}}, -\mathbf{B_{RZ}}]'$			
	W	SDF	Sharpe	W	SDF	Sharpe	W	SDF	Sharpe	
	p-value		ratio	p-value		ratio	p-value		ratio	
Panel A: FX styles based on all currencies										
		Bench.:	1.17		Bench.:	1.20			Bench.	+FXS
C+M	0.000	0.000	1.50	0.000	0.000	1.52	0.000	0.000	1.27	1.52
C+V	0.000	0.000	1.53	0.000	0.000	1.53	0.000	0.000	1.21	1.53
M+V	0.002	0.005	1.29	0.002	0.005	1.32	0.001	0.003	1.19	1.32
C+M+V	0.000	0.000	1.52	0.000	0.000	1.53	0.000	0.000	1.23	1.54
Panel B: FX styles based on G10 currencies										
		Bench.:	1.17		Bench.:	1.20			Bench.	+FXS
C+M	0.008	0.017	1.27	0.018	0.032	1.28	0.018	0.040	1.21	1.28
C+V	0.000	0.000	1.41	0.000	0.001	1.40	0.000	0.001	1.19	1.40
M+V	0.046	0.057	1.23	0.055	0.062	1.26	0.037	0.043	1.19	1.25
C+M+V	0.001	0.002	1.34	0.001	0.004	1.35	0.001	0.004	1.19	1.35

Kurzlebenslauf

Tim Alexander Kroencke

Ausbildung und akademischer Werdegang

2009 - 2013	ZEW Mannheim und Universität Mannheim
	Wissenschaftlicher Mitarbeiter im Forschungsbereich
	Internationale Finanzmärkte und Finanzmanagement
	Doktorand an der Fakultät für Betriebswirtschftslehre
2005 - 2009	Universität Freiburg
	Diplomstudiengang, Abschluss als Diplom-Volkswirt
2004 - 2005	Zivildienst
2002 - 2004	SEB Bank AG
	Ausbildung zum Bankkaufmann, Abschluss als Bankkaufmann (IHK)