Expected Returns in the Time Series and Cross Section: Empirical Evidence on Multifactor Asset Pricing Models and their Applications

Dissertation

zur Erlangung des akademischen Grades

Dr. rer. nat.

eingereicht an der

Mathematisch-Naturwissenschaftlich-Technischen Fakultät

der Universität Augsburg

von

Fabian Tobias Lutzenberger

Augsburg, August 2014

1. Gutachter:	Prof. Dr. Andreas W. Rathgeber
2. Gutachter:	Prof. Dr. Marco Wilkens
Tag der mündlichen Prüfung:	27.10.2014

Für meine Eltern

Contents

List of Abbreviations

AC_1	first-order autocorrelation coefficient
APT	arbitrage pricing theory
AR	autoregressive
BG	big–growth portfolio
BL	big–loser portfolio
BN	big–neutral portfolio
BV	big–value portfolio
BW	big–winner portfolio
C	four-factor model of Carhart (1997)
CAPM	capital asset pricing model
CE	cost of equity capital
Ch.	chapter
CRSP	Center for Research in Security Prices
CV	unrestricted version of the ICAPM of Campbell and Vuolteenaho (2004)
DJIA	Dow Jones Industrial Average
EMH	efficient markets hypothesis
Eq.	equation
EURIBOR	Euro Interbank Offered Rate
FF3	three-factor model of Fama and French (1993)
FF5	five-factor model of Fama and French (1993)
FF97	Fama and French (1997)
FIBOR	Frankfurt Interbank Offered Rate
GDP	gross domestic product
GLS	generalized least squares
GM09	Gregory and Michou (2009)
GMM	generalized method of moments
GRS	Gibbons, Ross, and Shanken (1989)
HL	ICAPM of Hahn and Lee (2006)
i.i.d.	independent and identically distributed
ICAPM	Intertemporal CAPM
IS	in-sample
KLVN	three-factor model of Koijen et al. (2010)
LIQ10	10 stock portfolios sorted by liquidity beta
max.	maximum
min.	minimum
MSC	Maio and Santa-Clara (2012)
OLS	ordinary least squares
OOS	out-of-sample
P	ICAPM of Petkova (2006)
PS	four-factor model of Pástor and Stambaugh (2003)
R^2	coefficient of determination
RRA	relative risk aversion of the representative investor
S&P	Standard & Poor’s
SBM25	25 stock portfolios sorted by size and book-to-market
Sec.	section
SG	small–growth portfolio
SL	small–loser portfolio
SM25	25 stock portfolios sorted by size and momentum

SN	small–neutral portfolio
std.	standard deviation
SV	small–value portfolio
SW	small–winner portfolio
t-mean	t-statistic of the mean
UK	United Kingdom
US	United States
VAR	vector autoregressive
vs.	versus
WLS	weighted least squares

List of Papers

This thesis is based on three papers I wrote as a doctoral student at the University of Augsburg. Each examines one or two of the five research questions considered in this thesis:

Research Question 1	Lutzenberger, Fabian T., 2014a, The predictability of aggregate returns on commodity futures, *Review of Financial Economics* 23, 120-130.
Research Question 2 and 3	Lutzenberger, Fabian T., 2014b, Multifactor models and their consistency with the ICAPM: Evidence from the European stock market, Discussion Paper WI-430, University of Augsburg, *European Financial Management,* forthcoming.
Research Question 4 and 5	Lutzenberger, Fabian T., 2014c, Industry cost of equity capital: European evidence for multifactor models, Discussion Paper WI-434, University of Augsburg.

I explicitly state that many paragraphs of this thesis are taken directly from these papers. Further papers I contributed to during this time that are related to but not directly incorporated in this thesis are as follows:

Isakovic, Vasko, Fabian T. Lutzenberger, and Stefan Stöckl, 2012, The multiperiod CAPM and the cross-section of asset prices, Discussion Paper WI-392, University of Augsburg, presented at the 25th European Conference on Operational Research, Vilnius, Lithuania.

Lutzenberger, Fabian T., Mikhail Patsev, and Stefan Stöckl, 2013, Asset valuation in a (realistic) conditional CAPM setting, Discussion Paper WI-437, University of Augsburg, presented at the 21st Annual Conference of the Multinational Finance Society, Prague, Czech Republic.

Lutzenberger, Fabian T., 2014, Are multifactor models consistent with the ICAPM? Further evidence from the US, Discussion Paper WI-454, University of Augsburg.

1. Introduction

Do expected asset returns vary through time?[1] Why do some assets exhibit higher average returns than others? How can factors that drive expected returns in the time series be linked to factors that explain the cross-sectional dispersion in average returns? How do these findings affect applications? These questions are so essential that Eugene Fama, Lars Peter Hansen, and Robert Shiller received the 2013 Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel for their empirical methods and applied work aimed at answering them (Royal Swedish Academy of Sciences, 2013). This thesis seeks to provide small but important contributions to these questions by investigating the time-series predictability of commodity futures returns through various factors, by testing various multifactor asset pricing models in the cross section of European stocks, and by examining whether these models are qualified to estimate the cost of equity capital (CE) of European industries.

The asset pricing literature has undergone two great revolutions (Cochrane, 2005, Ch. 20). The first revolution introduced the portfolio theory of Markowitz (1952, 1959), the capital asset pricing model (CAPM) of Sharpe (1964) and Lintner (1965a), and the efficient markets hypothesis (EMH) of Fama (1970a). The view that emerged through these papers is that asset returns are barely predictable from past returns or other factors, that expected returns are nearly constant through time, and that the behavior of asset prices is close to that of random walks; that investors only regard the means and (co-)variances of one-period asset returns to build efficient portfolios; and that variation in expected returns across assets is completely explained by differences in an asset's sensitivity to the market return (its market beta).

The second revolution comprises a variety of mainly empirical studies that changed this view substantially. First, since the 1980s, an accumulation of empirical evidence indicates that expected aggregate asset returns (e.g., the returns on broad stock market indices) are predictable by various factors and thus vary through time. These factors are past returns (e.g., Fama and French, 1988a), financial ratios such as the market dividend–price ratio (e.g., Campbell and Shiller, 1988a; Fama and French, 1988b), term structure variables such as the term spread (e.g., Fama and French, 1989), macroeconomic quantity variables such as the aggregate consumption–wealth ratio (Lettau and Ludvigson, 2001a), as well as several corporate decision variables such as the relation of equity issues to total new equity and debt issues (Baker and Wurgler, 2000).

Second, at the same time, empirical work reveals at least three "puzzles" in the cross section of expected asset returns that the CAPM seems to be unable to explain: (1) Small stocks (with low market capitalization) seem to have higher average returns than big stocks (Banz, 1981), (2) value stocks (with a high ratio of book value to price) tend to have higher average returns than growth stocks (e.g., Fama and French, 1992), and (3) past winners (with high returns over the past year) seem to outperform past losers (Jegadeesh and Titman, 1993). Fama and French (1993) show that their three-factor model (FF3) does a fairly good job of capturing empirical observations (1) and (2). Their model includes a size factor, constructed as the

[1] This chapter is based on the introductions of Lutzenberger (2014a, 2014b, 2014c).

difference between the returns on diversified portfolios of small and big stocks, as well as a value factor that represents the difference between the returns on diversified portfolios of value and growth stocks. Carhart (1997) augments this model with a factor that is constructed as the difference between the returns on diversified portfolios of stocks that have done well over the past year and the losers of the past year. The resulting four-factor model (C) seems to also work well in explaining the momentum anomaly, (3). A shortcoming of these two models is that the addition of their three factors to the market is essentially empirically motivated, whereas their theoretical foundation and economic motivation is somewhat weak.

Merton (1973) develops an Intertemporal CAPM (ICAPM) that links factors that forecast aggregate returns in the time series to factors that explain the cross-sectional variation of average returns. The model postulates that innovations to state variables that drive expected aggregate returns in the time series should show up as factors in the cross section, so that multiple factors are required to explain the cross-sectional variation in average asset returns. Merton assumes time-varying consumption–investment opportunities and investors who hedge themselves against these variations and thus weakens one of the static CAPM's main restrictive assumptions. The ICAPM serves as a potential theoretical explanation for the success of empirically motivated multifactor models such as FF3 and C. Beyond that, Campbell and Vuolteenaho (2004), Hahn and Lee (2006), and Petkova (2006), among others, develop the multifactor models CV, HL, and P, respectively, that they explicitly justify as empirical applications of the ICAPM. These models comprise innovations to a set of variables that seem to predict future stock market returns, such as innovation to the market dividend–price ratio, and seem to do well in explaining the CAPM anomalies, (1) to (3), pointed out earlier. In related work, Pástor and Stambaugh (2003), for example, propose a model, PS, that considers liquidity risk a determinant of average returns.

Maio and Santa-Clara (2012) emphasize that the ICAPM restricts the time-series and cross-sectional behavior of these state variables and factors. In particular, the candidate state variables must predict the distribution of aggregate returns, the market (covariance) price of risk must be economically plausible as an estimate of the representative investor's relative risk aversion, and the signs of the state variables' predictive slopes and factor risk prices must be consistent. The authors test whether the multifactor models FF3, the Fama–French (1993) five-factor model (FF5), C, PS, CV, HL, P, and the three-factor model of Koijen et al. (2010, KLVN) meet these ICAPM criteria within the US stock market. They conclude, however, that only FF3 and C can be justified as empirical applications of the ICAPM.

There are three ways to interpret the empirical findings that drive the second revolution in asset pricing, that is, time-varying expected returns and multiple factors in the cross section of expected returns, as done by, for instance, Fama and French (1996), based on the success of FF3. First, one may accept the EMH, that is, one may accept the hypothesis that security prices fully reflect all available information. Then, since asset pricing is concerned with how to measure risk and with the relation between risk and expected return, time-varying expected aggregate returns can be interpreted as a reflection of variation in risk aversion and aggregate (market) risk over time. Similarly, variation in average returns across assets that cannot be captured by the CAPM indicates that these assets differ from each other in a form of risk that is beyond the market beta. With regard to the three Nobel Prize recipients of 2013, this view or argumentation is mainly supported by Eugene Fama. Second, one may not accept the EMH. Under this assumption, variations in expected returns through time and average returns that cannot be fully explained by the CAPM can be interpreted as a result of the irrational behavior of investors and mispricing. Robert Shiller can be regarded as a representative of such argumentation.

Third, one may argue that these empirical findings are spurious and not real. In particular, one may argue that they are identified with data-snooping activities, that they reflect sample-specific characteristics that are just a result of random chance, or that they result from the use of inappropriate statistical tests that employ only small samples with few data points. This third interpretation can rely on, among other things, the fact that most empirical studies on expected returns in the time series and cross section employ solely US stock market data, while evidence from other stock markets and other asset classes is relatively sparse, as emphasized by various papers (e.g., Subrahmanyam, 2010; Artmann et al., 2012a, 2012b; Asness et al., 2013). The great impact of US stock market data is reflected by the fact that US data from one single database, that of the Center for Research in Security Prices, "provide[s] the foundation of at least one-third of all empirical research in finance over the past 40 years" (Economist, 2010).

While the first two interpretations appear difficult to test because of the joint hypothesis problem pointed out by Fama (1970a, 1991), it looks as if there are several ways to investigate the third interpretation. One way is to test whether the results obtained from US stock market data also hold for other stock markets and asset classes. Such out-of-sample tests have been called for by Fama (1991) and many others. This thesis is concerned with testing the third interpretation. In particular, it conducts an investigation on the commodity futures and European stock market data of (multi-)factor models that have been previously tested using US data in an attempt to assess their out-of-sample validity.

Beyond the idea of testing whether the results from US stocks are spurious or robust across different samples (i.e., a test of the third interpretation), extensions of studies based on US stocks to other asset classes and other stock markets help to reveal differences in the return-generating processes of different asset classes and stocks from different markets. Such differences might come from dissimilar asset characteristics, institutional settings, and investor preferences. Which (multi-)factor asset pricing model should be used for which asset class or market? This question is important in regard to the application of asset pricing models in different markets. For instance, European decision makers and regulation authorities might have to apply a different factor model for CE estimation than US decision makers.

In his Presidential Address to the American Finance Association, Cochrane (2011) points out that the asset pricing literature's second revolution will have a large impact on practical applications of asset pricing models such as portfolio management, performance evaluation, and corporate finance, accounting, and regulation, which still commonly rely on the random walk and CAPM view of asset pricing theory. For instance, many textbooks on corporate finance now recognize the CAPM's failures and advocate the use of alternative approaches to estimate CEs. As an example, Pratt and Grabowski (2010, p. 230) "conclude that beta alone does not fully measure the risk of most securities … [and] recommend that analysts use other risk measures beyond just beta.” Among a variety of other methods, they describe FF3. However, financial managers typically find no answers from the corporate finance literature regarding which alternative model to use in the end.

There seem to be at least two reasons to replace the CAPM with FF3 or C in regard to CE estimation. First, these two models seem to outperform the CAPM in explaining the cross section of expected stock returns and appear to be the current “industry standard” in the asset pricing literature (e.g., Artmann et al., 2012a). Second, the results of Maio and Santa-Clara (2012) indicate that the ICAPM may serve as a potential theoretical explanation for their empirical success. However, there are also at least three obstacles that are associated with applying these two models. First, their three additional factors are not any well-known economic variables, such as a return on a stock market index, but purely empirical constructs, what makes them difficult to “sell” to decision makers and regulation authorities (e.g., Cochrane, 2005, Ch. 20.2 with respect to C). Second, neither model directly builds on any rational theory of market equilibrium, such that there is still no consensus about whether the authors’ empirical success can be explained by rational investors who require compensation for risk (i.e., by the ICAPM) or by irrational investors who misprice stocks. Models of mispricing are regarded as inappropriate for setting CEs (e.g., Subrahmanyam, 2013). Third, Fama and French (1997) find that the CE estimates for US industries obtained from both the CAPM and FF3 are alarmingly imprecise. Gregory and Michou (2009) confirm these results for UK industries. They additionally examine C, which does, however, not seem to provide more accurate CE estimates.

An alternative could be multifactor asset pricing models that consider liquidity risk or which are explicitly justified by their authors as empirical applications of the ICAPM, such as PS, CV, HL, and P. These models might be superior to FF3 and C in estimating CEs, for several reasons: First, they typically perform as well as or better than these two models in explaining the cross-sectional variation in average stock returns. Second, they include mainly economically motivated factors, such as liquidity-related variables and innovations to economic state variables, rather than purely empirically-based factors. Such factors might be more sellable to decision makers and regulation authorities. Third, one might argue that models that are explicitly justified by their authors as empirical ICAPM applications have a stronger theoretical foundation.[2]

There is, however, no study that investigates whether multifactor asset pricing models that include economically motivated factors are more qualified than the CAPM, FF3, and C to produce accurate CEs. This thesis seeks to fill this research gap to some extent by comparing CEs for European industries that are obtained from the multifactor models PS, CV, HL, P, FF5, and KLVN, with CEs obtained from the CAPM, FF3, and C.

To be specific, this thesis examines five research questions. The first research question is concerned with testing the robustness of the conclusion that expected aggregate asset returns

[2] Maio and Santa-Clara’s (2012) results, however, counter this third argument. Additionally, the factors themselves that these models utilize are not derived from a theoretical model of economic equilibrium, albeit they are economically motivated.

are predictable by various factors and thus vary through time—a conclusion previously drawn on the basis of US stocks—by testing the time-series predictability of aggregate returns on an asset class other than US stocks, namely, commodity futures. For this purpose, I do not empirically test a single specific theory of commodity futures returns and its implications, such as the theory of storage of Kaldor (1939) and others. Instead, I test a large set of potential predictors. Most of these candidate predictors are standard choices in papers that study the time variation of expected aggregate returns on US stocks. Additionally, I suggest new factors, most of which are commodity specific. In particular, I conduct predictive regressions that use the future return on an equally weighted portfolio of 27 commodity futures (without the return on collateral) as the response variable. The right-hand sides of these regressions comprise the current values of the subsets of 32 potential predictors from the stock market, bond market, macroeconomics, and commodity market. The main sample period is from January 1972 to June 2010, with a monthly sample frequency. Predictive regressions are the most common approach for forecasting aggregate returns (Kelly and Pruitt, 2013). If the returns are unpredictable, regression coefficients beyond a constant should be insignificant in such models and these models should not provide forecasts of future returns that are more accurate than the historical average of past returns. I evaluate both in-sample and out-of-sample predictability. For the in-sample analysis, I employ single long-horizon predictive regressions with horizons of one, three, 12, 24, 36, 48, and 60 months (e.g., Maio and Santa-Clara, 2012), as well as a procedure that selects the "best" multiple-variable regression out of the variety of candidate predictors I consider, as proposed by Bossaerts and Hillion (1999) and Zakamulin (2013). The out-of-sample evaluation comprises forecasts from single predictive regressions from the model selection procedure following Bossaerts and Hillion and Zakamulin and from several combination forecasts proposed by Rapach et al. (2010).

The second and third research questions motivate a European investigation of the eight multifactor asset pricing models consisting of FF3, C, FF5, PS, CV, HL, P, and KLVN in an attempt to test their ability to explain the cross section of average returns and their consistency with the ICAPM out of sample. For this purpose, I extend Maio and Santa-Clara's (2012) US study to a large sample of stocks from 16 European countries. According to Fama and French (2012), these countries account for, on average, 30% of global market capitalization and represent the second largest integrated stock market region (after North America). Five of the eight multifactor models investigated, that is, PS, CV, HL, P, and KLVN, have never had their explanatory power tested for European stock returns. Moreover, none of the eight multifactor models has ever been tested for consistency with the ICAPM using data from outside the US.

Finally, the fourth and fifth research questions are not concerned with the empirical performance or validity of factor models but, give attention to their applicability or usefulness for practical purposes, specifically with CE estimation. In particular, I examine whether the CAPM and the multifactor models FF3, C, FF5, PS, CV, HL, P, and KLVN are qualified to estimate CEs for European industries. Moreover, I investigate whether the CEs obtained from these nine factor models—estimated using different econometric techniques—differ from each other, to examine whether the choice of the factor model and estimation technique is at all important for CE estimation. With these research questions, I extend the work of Fama and French (1997) and Gregory and Michou (2009) in two ways. First, I provide a robustness check of these two studies by examining whether CEs obtained from the CAPM, FF3, and C are also inaccurate for European industries. Second, I investigate whether multifactor models that include economically motivated factors provide more accurate CEs than the CAPM, FF3, and C, whose factors in addition to the market return are solely empirically-based.

The remainder of this thesis is organized as follows. Chapter 2 reviews the literature on expected asset returns in the time series and cross section and introduces the research questions. Chapter 3 presents the data and variables and Ch. 4 presents the empirical methodology. Chapter 5 presents the empirical results, which are discussed in Ch. 6. Chapter 7 provides a summary and concludes the thesis.

2. Theory, Prior Evidence, and Research Questions

The literature on expected assets returns is vast and it is impossible to completely cover it. Hence, I focus on work that is, in my opinion, mandatory for understanding the motivations and implications of the empirical tests that are conducted in this thesis. Moreover, the focal points are the economic intuitions and implications behind asset pricing models and the empirical evidence that is associated with them. Consequently, I do not focus on the technical aspects of these models, such as their mathematical derivations. Moreover, I do not present the large amount of literature that is mainly concerned with the methodologies and techniques to empirically test asset pricing models. Instead, I concentrate on the empirical facts that are obtained using these methods and present the econometric techniques that are applied in Ch. 4 of this thesis.

Moreover, while most asset pricing theories consider assets in general, empirical evidence predominantly focuses on US stocks. The empirical examinations in this book provide evidence from another asset class, namely, commodity futures and from the European stock market. For this reason, I start by presenting theories together with empirical evidence for US stocks and later highlight the rarer evidence on other asset classes and markets outside the US in separate sections.

I start with the literature on expected returns in the time series. Other surveys on this topic, which are more extensive in some ways, have been conducted by Fama (1991), Welch and Goyal (2008), Lettau and Ludvigson (2010), Koijen and van Nieuwerburgh (2011), the Royal Swedish Academy of Sciences (2013), and Campbell (2014), among others. I then outline the literature on expected returns in the cross section. The interested reader can refer to Campbell (2000), Fama and French (2004), Subrahmanyam (2010), Daniel and Titman (2012), and Goyal (2012) for more elaborate presentations with regard to certain aspects. Jagannathan et al. (2010a, 2010b) and Lewellen et al. (2010), among others, provide valuable reviews that are mainly concerned with the empirical methodologies for testing asset pricing models.

2.1. Expected Returns in the Time Series (Time-Series Predictability)

2.1.1. Stock returns in the US

The consensus conclusion of the early (pre-1970) literature on the predictability of returns, such as the works of Fama (1965) and Fama and Blume (1966), is that stock returns are nearly unpredictable from past returns or other variables, expected returns are almost constant through time, and stock prices behave similarly to random walks. These early studies typically focus on the predictability of short-term returns (Fama, 1991; Cochrane, 2005, Ch. 20). For instance, Fama (1965) examines the autocorrelation coefficients of the daily as well as four-, nine-, and 16-day returns of the 30 Dow Jones Industrial Average (DJIA) stocks over the sample period 1956 to 1962, using one to 10 lags. A stock's first-order autocorrelation (one lag), for example, can be estimated by regressing its return at time $t+1$, R_{t+1}, on its return at time t:

$$R_{t+1} = a + bR_t + u_{t+1}, \tag{2.1}$$

where a is a constant and u_{t+1} denotes a disturbance term with zero conditional mean, $E_t(u_{t+1}) = 0$. It follows that the conditional expectation at time t (the expected return conditional on the information set at time t) is given by

$$E_t(R_{t+1}) = a + bR_t. \tag{2.2}$$

The slope coefficient (the first-order autocorrelation), b, can be interpreted as a measure of how important R_t is in predicting R_{t+1}. If the estimate of b is not significantly different from zero and the conditional expected return is solely described by a, returns are unpredictable from past returns and expected returns are constant through time (according to the model). Regarding the first-order autocorrelations of daily returns, Fama (1965) finds that 23 are positive and 11 are significantly different from zero. However, their magnitudes are very small in comparison to the variation of daily returns, such that Fama concludes "that dependence in successive price changes is either extremely slight or completely non-existent" (p. 74).

The early studies on return predictability are followed by many papers that use—since time flies—longer time series and more sophisticated econometric techniques. For instance, Conrad and Kaul (1988) and Lo and MacKinlay (1988) examine the autocorrelations of weekly returns on portfolios of NYSE and AMEX stocks sorted by size (market capitalization, price times shares outstanding). They find evidence of stronger and more significant autocorrelation in returns than former evidence on individual stocks suggests; that is, past returns seem to explain a higher proportion of the variation in realized returns, which mainly results from the diversification effects of portfolios. However, the authors also show that predictability is considerably higher for small stocks. Since the trading of small stocks is usually associated with higher transactions costs, the predictability seems to be difficult to exploit. Nevertheless, evidence from the 1980s changes the 1970s view to some extent. Overall, these studies show that short-term returns are somewhat predictable from past returns and that this effect is statistically significant, as emphasized by Fama (1991).

Researchers in the 1980s started to also examine autocorrelations in long-horizon returns. Fama and French (1988a) show that the autocorrelations of returns on portfolios of NYSE stocks rise with horizons (in absolute terms). They are close to zero for short-horizon returns, but reach values around -0.25 to -0.4 for three- to five-year returns. Poterba and Summers (1988) support this evidence by comparing the variance of short-horizon returns to that of long-horizon returns. Nevertheless, Fama (1991) states that these tests still provide only weak evidence against the null hypothesis of no autocorrelation in returns, that is, there is no predictability of future returns from past returns. In particular, realized returns seem to be noisy instruments for measuring conditional expected returns. On the basis of this finding, researchers began to extensively search for economic variables that are better predictors of future returns.

Evidence that returns are correlated with other economic variables was already found in the 1970s. For instance, Bodie (1976) and Fama (1981) as well as Fama and Schwert (1977) show that monthly stock returns are negatively related to (expected) inflation and short-term interest rates, respectively. However, the predicted variation in expected returns is still a small part of the variance of realized returns (Fama, 1991). Similarly, Rozeff (1984) and Shiller (1984) provide evidence that the aggregate dividend–price ratio predicts short-horizon returns.

Fama and French (1988b) examine the power of the aggregate dividend–price ratio to forecast one- and four-month returns as well as one-, two-, three-, and four-year returns on a value-weighted and an equal-weighted portfolio of NYSE stocks. Cochrane (2011) updates some of Fama and French's regressions using annual data from 1947 to 2009:

$$R^e_{t,t+q} = a_q + b_q DY_t + u_{t,t+q}, \tag{2.3}$$

where $R^e_{t,t+q}$ is the Center for Research in Security Prices (CRSP) value-weighted market return in excess of the three-month US Treasury bill return from time t to $t+q$, a_q is a constant, DY_t is the value of the market dividend–price ratio at time t, and b_q is the slope coefficient. Cochrane uses the forecasting horizons $q = 1$ year and $q = 5$ years. Like Fama and French (1988b), Cochrane computes the t-statistic of the slope coefficient associated with the five-year returns following Hansen and Hodrick (1980) to correct for the serial correlation in the residuals, which results from the overlapping return observations. In addition, $u_{t,t+q}$ is a forecasting error, which has a zero conditional mean. On this basis, the conditional expected return at time t is

$$E_t\left(R^e_{t,t+q}\right) = a_q + b_q DY_t. \tag{2.4}$$

Table 1 shows Cochrane's results, which are qualitatively similar to Fama and French's (1988b).

Table 1. Predictive regressions.
This table is from Cochrane (2011, Table I). It shows the results of q-horizon predictive regressions with the CRSP value-weighted percent return in excess of the three-month US Treasury bill return as the dependent variable and the percent aggregate dividend–price ratio as the independent variable. The regression employs annual data from 1947 to 2009. The second column displays the estimated slope coefficient, the third row the t-statistic of the slope coefficient computed using Hansen and Hodrick's (1980) standard errors, and the fourth column displays the R^2 of the regression. Moreover, the second to last column shows the standard deviation of the fitted value and the last column displays its proportion to the unconditional mean of the dependent variable.

Horizon q	b_q	$t(b_q)$	R^2	$\sigma\left[E_t\left(R^e_{t,t+q}\right)\right]$	$\frac{\sigma\left[E_t\left(R^e_{t,t+q}\right)\right]}{E\left(R^e_{t,t+q}\right)}$
1 year	3.8	(2.6)	0.09	5.46	0.76
5 years	20.6	(3.4)	0.28	29.3	0.62

First, note that the slope coefficient associated with the one-year horizon is significantly different from zero. Consequently, the equity premium seems to vary through time; it is not solely described by the regression's constant. Moreover, the R^2 value shows that 9% of the variation in realized returns one year ahead is explained by variation in DY_t, which does not appear impressive at first glance. However, Cochrane (2011) emphasizes that this result does, in fact, have huge economic significance. Regarding the magnitude of the slope coefficient, one can see that an increase in DY_t of one percentage point results in an increase in expected excess return of nearly four percentage points. Second, observe that the expected one-year return varies greatly through time. In particular, the last column shows that the variation through time in conditional expected returns is at least as high as the unconditional expected return. To be specific, the equity premium varies between -3.7% and 18.1% according to the rule of thumb of plus or minus two standard deviations. Third, note that the R^2 increases with the horizon. Variation in DY_t seems to explain as much as 28% of the variation of returns five years ahead. Cochrane (2005, Sec. 20.3) states that the long-horizon regression can be interpreted as a "magnifying glass": It makes a small fact, that is, slight predictability in short-horizon returns, economically interesting. The reason for the increasing R^2 is that DY_t is a slow-moving variable, as shown by its autocorrelation coefficient of around 0.9. However, several authors, such as Boudoukh et al. (2007b) and Hjalmarsson (2011), caution against concluding that predictability is stronger for long-horizon returns just because the slope coefficients and R^2 values rise. Instead, it seems that the effects that are already present in short-horizon returns simply become more obvious and economically interesting.

In related work, Campbell and Shiller (1988a) derive a linear approximation to an exact relation between prices, returns, and dividends, termed the approximate present value identity. The model can be used to examine to what fraction of variations in the aggregate dividend–price ratio are related to variations in expected dividends and expected returns. Cochrane (2011, p. 1050) provides updated evidence for this model and concludes "that all price–dividend ratio volatility corresponds to variation in expected returns. None corresponds to variation in expected dividend growth, and none to 'rational bubbles.'"

Cochrane (2011) emphasizes that the 1970s view, that is, the assumption of constant expected returns, would have suggested the opposite, that is, low dividends relative to prices represent an expectation of higher future dividend growth. Instead, however, low prices relative to dividends seem to entirely predict higher subsequent returns.

With Fama and French's (1988b) study, the 1970s view of constant expected returns seemed to be finally out-of-date. Following this early evidence, the asset pricing literature proposed a variety of further variables to forecast future returns. Koijen and van Nieuwerburgh (2011) categorize these variables into four groups: (1) financial ratios, such as the aggregate dividend–price ratio, (2) term structure variables, (3) macroeconomic quantity variables, and (4) corporate decision variables.

Among **financial ratios**, Campbell and Shiller (1988b) use different aggregate earnings–price ratios of the Standard & Poor's (S&P) 500 index to predict S&P 500 returns. These ratios employ either earnings lagged by one year or the 10- or 30-year moving averages of past earnings. Similarly, Kothari and Shanken (1997) propose the aggregate book-to-market ratio of the DJIA to forecast returns on the CRSP equal-weighted and value-weighted market portfolios over the period 1926 to 1991. In related work, Pontiff and Schall (1998) state that the predictive power of the DJIA book-to-market ratio is specific to the period 1926 to 1960, while the S&P 500 book-to-market ratio shows forecasting power over the period 1960 to 1994. Moreover, Lamont (1998) uses the aggregate dividend–payout ratio, that is, the ratio of dividends to earnings, to predict excess returns on the S&P 500. Furthermore, Campbell and Vuolteenaho (2004) propose the small-stock value spread, which is the difference between the log book-to-market ratios of a portfolio of small stocks with high book-to-market ratios and a portfolio of small stocks with low book-to-market ratios, as a predictor of CRSP value-weighted excess market returns. Beyond that, Guo (2006a) recommends a measure of aggregate stock market volatility, namely, the realized stock market variance, to forecast the equity premium. Polk et al. (2006) build on the central prediction of the capital asset pricing model (CAPM) that stocks' expected returns are linearly related to their betas and that the slope of this relation equals the excess return on the market portfolio. The authors use this prediction to construct a variable, the cross-sectional beta premium, that shows predictive power for CRSP value-weighted excess market returns. Baker and Wurgler (2006, 2007) construct a stock market sentiment index that is based on the first principal component of six US sentiment proxies: NYSE turnover, dividend premiums, closed-end fund discounts, the number of and first-day returns on initial public offerings, and equity share in new issues. Among other things, Baker and Wurgler (2006) use their sentiment index to predict the returns on a variety of characteristic-based long–short portfolios of US stocks. Boudoukh et al. (2007a) advocate the ratios of payout (dividends plus repurchases) to price as well as of net payout (dividends plus repurchases minus issuances) to price instead of the dividend–price ratio to forecast CRSP value-weighted excess market returns. Bollerslev et al. (2009) and Drechsler and Yaron (2010) advocate the variance risk premium, defined as the difference between implied and realized variation in aggregate stock market returns, as a predictor of future returns. The implied variation is obtained from "model-free" options implied volatilities using the new Chicago Board Options Exchange Volatility Index and realized variation is measured with high-frequency S&P 500 intraday returns. The authors show that high-volatility premiums forecast high future S&P 500 returns. Kelly and Pruitt (2013) extract a single factor from the cross section of disaggregated book-to-market ratios (i.e., of the book-to-market ratios of a variety of portfolios). They find that this factor shows substantial predictive power for CRSP value-weighted market returns. Li et al. (2013) propose the aggregate implied cost of capital as a predictor of future market returns. Jacquier and Okou (2014) separate realized variance into continuous volatility and jump components. They show that continuous volatility predicts medium- to long-term returns, while jumps show no predictive power over these returns.

Term structure variables that appear to show predictive power for future stock market returns are the US Treasury bill rate (Campbell, 1987); the yield and realized return on long-term US government bonds (Welch and Goyal, 2008); the term structure spread, which is the long-term yield on US government bonds minus the US Treasury bill rate (Campbell, 1987; Fama and French, 1989); the default yield spread, which represents the difference between the yields on US high-yield corporate bonds (medium or low credit quality) and low-yield corporate or government bonds (Keim and Stambaugh, 1986; Fama and French, 1989); the default return spread, which is similar to the default yield spread but in terms of realized returns instead of yields (Welch and Goyal, 2008); and the Cochrane–Piazzesi factor (Cochrane and Piazzesi, 2005).[3]

Macroeconomic quantity variables that are proposed as predictive variables for stock market returns are the ratio of aggregate investment to aggregate capital for the whole US economy (Cochrane, 1991); the aggregate consumption–wealth ratio, characterized as the residual from a cointegrating relation between US consumption, asset wealth, and income (Lettau and Ludvigson, 2001a); the ratio of US consumption to aggregate stock prices (Menzly et al., 2004); the ratio of US labor income to consumption (Santos and Veronesi, 2005); the price–output ratio, which is the ratio of aggregate stock prices to the gross domestic product (Rangvid, 2006); the housing–collateral ratio, defined as the ratio of aggregate housing wealth to human wealth in the US (Lustig and van Nieuwerburgh, 2005); the housing–non-housing consumption ratio (Piazzesi et al., 2007); and an aggregate US dividend–price ratio corrected for demographics, that is, for the ratio of middle-aged to young (Favero et al., 2011).

Corporate decision variables used to forecast stock market returns are the share of equity issues in total US new equity and debt issues (Baker and Wurgler, 2000), as well as net equity expansion, which is the ratio of net equity issues by NYSE stocks to the total market capitalization of NYSE stocks (Welch and Goyal, 2008).

In an influential study, Welch and Goyal (2008) comprehensively examine a large portion of the predictive variables proposed as of early 2006, using data up to 2005. Their study is motivated by the fact that different papers tend to examine different variables using different econometric techniques, sample periods, proxies for the equity premium, and so forth, so that the results are difficult to compare. Welch and Goyal examine each variable separately and employ a "kitchen sink" predictive regression that includes all variables at once, as well as a model selection procedure. The latter selects at each time t the "best" multiple-variable predictive regression (with regard to a pre-specified model selection criterion) out of the 2^K competing regressions that can be constructed from all possible combinations of the K variables under consideration. Such a procedure is advocated by Bossaerts and Hillion (1999). In contrast to most of the aforementioned studies, Welch and Goyal regard not only the in-sample (IS) predictive power of each variable, but also their out-of-sample (OOS) performances. An IS examination conducts regressions of returns on lagged predictors over the whole sample period. The predictive power of the variable of interest is then typically assessed by regarding the statistical significance of the estimated slope coefficient and the regression's R^2. In contrast, an OOS analysis conducts regressions at each point in time t, using data only up to t, to produce a forecast of the return at $t+1$. The $t+1$ return forecast is then compared to the realized return at $t+1$ as well as to a forecast of the $t+1$ return obtained from the average of the returns realized up to t. The latter is the best forecast for the $t+1$ return under the null hypothesis of constant expected returns. Hence, an OOS examination adopts the perspective of an investor who has used these variables to time the market and shows whether the investor would have been better off using the historical mean return instead. Welch and Goyal conclude "that a healthy skepticism is appropriate when it

[3] I present the latter in more detail in Ch. 2.1.2.

comes to predicting the equity premium" (2008, p. 1456). In particular, they state that most variables are no longer significant even IS and the significance of many variables that is found in earlier studies seems to be based largely on the 1973 to 1975 oil shock and stock market collapse. Similarly, most variables show poor OOS performance, so that these variables would have not helped an investor to time the market any better than the historical average return would have.

Welch and Goyal's (2008) study and its rather surprising results have met with critique and controversy and have motivated researchers to search for more sophisticated forecasting techniques in order to—after all—reject the null hypothesis of unpredictable stock returns. In particular, Cochrane (2008) criticizes Welch and Goyal's OOS analysis and states that it is not a better test of predictability than an IS examination, although it gives interesting insight into the practical usefulness of return forecasts. Moreover, Cochrane argues that the *absence* of dividend growth predictability provides much stronger evidence against the null hypothesis of constant expected returns than does the *presence* of return predictability in predictive regressions. The author concludes that the observed variation in the aggregate dividend–price ratio is definitely not due to dividend growth forecastability. Consequently, it must result from return predictability. Hence, the unpredictability of dividend growth seems to be "the dog that does not bark" (Cochrane, 2008, p. 1535) in the case against unpredictable returns. Moreover, Campbell and Thompson (2008) state that posing economically meaningful restrictions on the regression coefficients and return forecasts can result in useful OOS predictions. Rapach et al. (2010) provide further evidence in favor of OOS return forecastability. They argue that combining the forecasts from individual variables (i.e., from single-variable predictive regressions) provides OOS forecasts that are significantly better than the historical mean return. In particular, a combination of individual forecasts reduces forecast volatility and considers information from numerous predictive variables at once. Zhu and Zhu (2013) show that this combination approach can be improved by augmenting it with a regime-switching model. Ferreira and Santa-Clara (2011) divide aggregate stock returns into three components—the dividend–price ratio, earnings growth, and price-earnings ratio growth—and advocate their separate prediction. Like the combining approach of Rapach et al., their sum-of-the-parts method seems to significantly beat the historical average return in forecasting future returns OOS. In related work, Zakamulin (2013) provides evidence that the premium on small stocks is predictable both IS and OOS.[4] Even before Welch and Goyal's study, Avramov (2002) provides evidence that stock returns are predictable in the presence of model uncertainty, using Bayesian model averaging.

Altogether, it seems that evidence that aggregate returns on US stocks are predictable by various factors and that the expected aggregate returns thus vary through time predominate studies, such as that of Welch and Goyal (2008), that counter or criticize this conclusion. In particular, more recent studies, such as that of Rapach et al. (2010), that use longer sample periods and more sophisticated econometric techniques seem to speak strongly in favor of return predictability. Nevertheless, it seems to be essential to test the OOS validity of the

[4] The small-stock premium (size factor) is presented in Section 2.2.1.3.

predictive variables proposed to forecast US stock returns by investigating whether these variables also show predictive power for international stock returns and other asset classes.

2.1.2. International stock returns and other asset classes

Time-series predictability is also studied in international stock markets and other asset classes, not least to assess whether the results obtained from US stock returns are robust across different samples. Paye and Timmermann (2006) examine the predictive power of the aggregate dividend yield, the short-term interest rate, the term structure spread, and the default yield spread for aggregate excess returns in 10 countries, including all G7 members. They construct country-specific estimates for the first three predictors, but use US data for the default yield spread. They find common as well as country-specific structural breaks in the predictive regressions for most countries and the relationships between expected excess returns and these four predictors seem to change considerably after a break. Ang and Bekaert (2007) study the predictability of aggregate excess stock returns in the US, UK, Germany, and France by the aggregate dividend–price ratio and the short-term interest rate (they obtain country-specific estimates for both predictors). Among other things, they conclude that the short-term interest rate is a more robust predictor than the dividend–price ratio. Hjalmarsson (2010) studies the predictive power of country-specific proxies for the dividend–price ratio, the earnings–price ratio, the short-term interest rate, and the term structure spread in 40 international markets. The author finds that the forecasting abilities of the short-term interest rate and the term structure spread are quite robust within developed markets, while the earnings–price ratio and the dividend–price ratio do not seem to be robust predictors of international (excess) stock returns, confirming the results of Ang and Bekaert. Further international evidence is provided by Ferson and Harvey (1993), Harvey (1995), Su et al. (2009), and others.

Predictability is also studied in asset classes other than stocks. Moskowitz et al. (2012) find positive autocorrelations, which they term time-series momentum, in the returns of a variety of futures and forward contracts associated with country equity indexes, currencies, commodities, and sovereign bonds. In particular, a security's past 12-month excess return seems to be a positive predictor of its future return. Likewise, Cochrane (2011, p. 1051) emphasizes that the finding that the dividend–price ratio does not predict future dividend growth but future excess stock returns is pervasive across markets and

> For stocks, bonds, credit spreads, foreign exchange, sovereign debt, and houses, a yield or valuation ratio translates one-for-one to expected excess returns, and does not forecast the cashflow or price change we may have expected. In each case our view of the facts has changed completely since the 1970s.

In particular, one-year returns on one- to five-year US government bonds seem to be predictable by one-year forward rates on these bonds (Fama and Bliss, 1987) and a high value of the term structure spread seems to indicate that the one-year returns on long-term government bonds are better than those of short-term bonds (Campbell and Shiller, 1991). Cochrane and Piazzesi (2005) advocate a linear combination of forward rates, termed the Cochrane–Piazzesi factor in later studies, to predict excess bond returns. To be specific, they obtain their predictive factor as the fitted value from the following regression of an average of excess bond returns on forward rates:

$$\frac{1}{4}\sum_{n=2}^{5} rx_{t+1}^{(n)} = \gamma_0 + \gamma_1 y_t^{(1)} + \gamma_2 F_t^{(2)} + \cdots + \gamma_5 F_t^{(5)} + \bar{\varepsilon}_{t+1}, \tag{2.5}$$

where $rx_{t+1}^{(n)}$ is the excess log return (over the yield of a one-year discount bond at time t) from buying an n-year discount bond at time t and selling this bond as an $(n-1)$-year bond at time $t+1$, $y_t^{(1)}$ is the log yield of a one-year bond at time t, and $F_t^{(n)}$ is the log forward rate at time t for loans between time $t+n-1$ and $t+n$. Moreover, γ_0 is the intercept, γ_n is a slope coefficient, and $\bar{\varepsilon}_{t+1}$ is a disturbance term.

Similar to government bonds, returns on private-issuer money market securities appear to be predictable by term and default premiums (Fama, 1986) and excess returns on federal funds futures contracts seem to be forecastable by macroeconomic quantity variables such as employment growth, as well as by term structure variables such as the term structure spread and the default yield spread (Piazzesi and Swanson, 2008).

Moreover, returns on currencies appear to be predictable by international interest rate spreads and other macroeconomic quantity and term structure variables (Hansen and Hodrick, 1980; Fama, 1984; Piazzesi and Swanson, 2008), as well as by dividend yields and past excess stock returns (Bekaert and Hodrick, 1992). In related work, Bakshi and Panayotov (2013) provide evidence that the payoffs of currency carry trade strategies are forecastable by changes in a commodity price index, changes in average currency volatility, and a variable that measures global liquidity. Recently, Lustig et al. (2014) advocate two variables to predict returns on the US dollar: US industrial production growth and the average forward discount on non-US currency against the US dollar, which represents the difference between the average short-term interest rate of non-US developed countries and the US short-term interest rate.

Furthermore, Gourinchas and Rey (2007) find that returns on the US net foreign portfolio (the return on foreign assets relative to US assets) is predictable by a measure of external imbalances. Similarly, Cochrane (2011) provides evidence that the return on a US house price index is predictable by the aggregate ratio of house prices to rents. In particular, a high ratio of house prices to rents predicts low subsequent housing returns. Similar evidence is provided by Campbell et al. (2009) for residential real estate and by Plazzi et al. (2010) for commercial real estate.

Finally, there is some evidence that returns on commodity (futures) are predictable. Fama and French (1987) test two theories of commodity futures prices. The first one is the theory of storage of Kaldor (1939), Working (1948, 1949), Brennan (1958), and Telser (1958), which relates the difference between contemporaneous commodity spot and futures prices to foregone interest (because of the commodity being stored), warehousing costs, and a convenience yield on inventory. Fama and French show empirically that interest rates as well as seasonals in convenience yields predict variation in spot prices. The second model divides the futures price into an expected risk premium and a prediction of the maturity spot price (e.g., Keynes, 1930; Hicks, 1939; Cootner, 1960; Dusak, 1973). Fama and French show that 10 of 21 commodity futures exhibit forecast power for the future spot price and five commodities have time-varying expected risk premiums. Similarly, Narayan et al. (2013) examine four commodities (oil, gold, silver, and platinum) and find that current futures returns forecast future spot returns, that is, the commodity futures market predicts the commodity spot market. Moreover, the authors provide evidence that, except for platinum, this predictability is exploitable by technical trading rules. More evidence that the commodity futures market predicts the commodity spot market is provided by Coppola (2008) with regard to oil, among others. Further predictors of commodity (futures) returns are hedging pressure, measured by the net positions of hedgers in futures markets (e.g., De Roon et al., 2000); various macroeconomic quantity and term structure variables such as bond yields, inflation rates, the term spread, and the default spread (Erb and Harvey, 2006, and articles cited therein); capital flows into commodity markets, measured by open interest (Hong and Yogo, 2012); physical inventory levels (Gorton et al., 2013); and the futures percentage basis, also known as the futures (cost of) carry (Szymanowska et al., 2014). The latter is interpreted by Szymanowska et al. (2014) as the commodity analogy of a return-predicting valuation ratio, like the dividend–price ratio is for aggregate stock returns. Consequently, the "pervasive phenomenon" of return-predicting yields or valuation ratios (Cochrane, 2011, p. 1051) seems to also apply to commodity markets. Finally, in related research, Jensen et al. (2000, 2002) show that a measure of the US monetary policy significantly predicts the performance and role of commodity futures in mean–variance-efficient portfolios.

2.1.3. Research question on expected returns on commodity futures in the time series

The previous two sections present much empirical evidence that suggests that aggregate asset returns are predictable in the time series by various factors and that expected returns vary through time. The vast majority of this evidence is based solely on US stocks, whereas evidence from international stock markets and other asset classes appears to be relatively sparse. The latter is, however, important to test for the robustness of the results obtained from US stock market data. Beyond that, revealing the commonalities and differences across markets and asset classes in regard to the predictive abilities of different variables is important for the application of these models.

One asset class that offers only little evidence of its time-series predictability is commodity futures. In particular, the studies mentioned in Section 2.1.2, such as those of Fama and French (1987) and Szymanowska et al. (2014), empirically test either specific theories or only a few potential predictors of commodity returns. However, no study comprehensively examines a large set of potential predictors as Welch and Goyal (2008) does for US stocks. Since all commodity studies mentioned earlier employ different data and sample periods, their results are difficult to compare. Moreover, there is a variety of financial ratios, term structure variables, macroeconomic quantity variables, and corporate decision variables whose predictive powers for US stock returns have been extensively tested, but whose forecasting powers for commodity futures returns have not yet been investigated (e.g., the aggregate consumption–wealth ratio of Lettau and Ludvigson, 2001a). Furthermore, there are a variety of stock-specific variables, such as the realized stock market variance proposed by Guo (2006a), whose abilities to predict US stock returns have been tested, but whose commodity analogies have not yet been proposed as forecasters of commodity returns. Finally, there is no study that tests the OOS predictability of commodity returns.

The first research question of this thesis deals with this research gap by examining the time-series predictability of commodity futures returns. The research question is as follows.

RESEARCH QUESTION 1: Are aggregate returns on commodity futures predictable?

I contribute to the literature on the predictability of commodity returns by following an empirical asset pricing approach as described in Fama and French (2013); that is, by working backward, I come from the empirical side and test a large set of potential predictors. Most of these candidate predictors are standard choices in studies of the return predictability of US stocks. In addition, I propose new factors, most of which are commodity-specific analogies to variables that seem to forecast US stock returns. For instance, I construct and test a commodity market variance variable as an analogy to Guo's (2006a) stock market variance. Finally, I test not only the IS predictive powers of these factors, but also their OOS forecasting abilities.

2.2. Expected Returns in the Cross Section (Cross-Sectional Predictability)

I choose a mainly chronological order to present the literature on expected returns in the cross section. One alternate possibility is to present all the theory and then present the empirical evidence afterward. However, theories are inspired by empirical work and vice versa, so I refrain from dividing them. As in the overview of time-series predictability, I focus on work that is mandatory for the empirical tests in this thesis. For this reason, I do not present the arbitrage pricing theory (APT) introduced by Ross (1976) and its empirical applications, consumption-based models such as that of Breeden (1979), as well as conditional versions of the CAPM, as presented by Jagannathan and Wang (1996), although these are important cross-sectional asset pricing models.

2.2.1. Stock returns in the US

2.2.1.1. The CAPM and its early empirical tests

The CAPM of Sharpe (1964) and Lintner (1965a) may be called the most famous and most widely used asset pricing model (e.g., Cochrane, 2005, Sec. 9.1). The CAPM relates the expected return in excess of the risk-free rate on any asset i to its market beta:

$$E(R_i) - R_f = \beta_{im}\left[E(R_m) - R_f\right], \tag{2.6}$$

where R_i is the return on asset i, R_f is the risk-free interest rate, R_m is the return on the market portfolio, and β_{im} is the market beta of asset i, which is the covariance between the return on asset i and the market return divided by the variance of the market return:

$$\beta_{im} = \frac{Cov(R_i, R_m)}{Var(R_m)}. \tag{2.7}$$

The market beta represents the sensitivity of the return on asset i to variation in the return on the market portfolio and serves as the measure of the risk of asset i. For this reason, it is also called the risk loading on the market return. The term $E(R_m) - R_f$ represents the risk premium on the return on any asset per unit of market beta and is commonly called the market beta risk price, the market risk premium, or the equity premium. Hence, the model states that an asset that covaries positively with the return on the market (that has a positive market beta) earns a risk premium over the risk-free rate. An economic intuition for this postulation is as follows (e.g., Maio and Santa-Clara, 2012; MSC hereafter): An asset that shows a positive market beta produces a high payoff in times of high market returns (high returns on aggregate wealth), but does poorly in times of low returns on the market as a whole. Hence, such an asset does not provide the investor with a hedge against changes in current wealth, so that a risk-averse investor requires this asset to offer a premium over the risk-free rate in order to invest in it.

Sharpe (1964) and Lintner (1965a) derive the CAPM by adding two key assumptions to Markowitz's (1952, 1959) portfolio selection model: *complete agreement*, that is, investors agree on the joint distribution of asset returns from t to $t+1$ (which is the true one), as well as *borrowing and lending at a risk-free rate*, which means that all investors have the opportunity to borrow or lend at the same risk-free rate, which is independent of the amount borrowed or lent (Fama and French, 2004). In its original form, the CAPM is, as Markowitz's portfolio theory is, a two-period model, that is, the model considers only two points in time, t and $t+1$. Beyond that, the pricing equation stated in Eq. (2.6) can be derived from several other sets of modeling assumptions that may include more than two periods. For instance, Cochrane (2005, Sec. 9.1) shows how to derive the CAPM on the basis of a consumption-based model under (1) two periods and a quadratic utility function, (2) two periods, exponential utility, and normally distributed asset returns, (3) an infinite horizon, quadratic utility, and asset returns that are independent and identically distributed (i.i.d.) over time (i.e., the investment opportunity set does not change through time), and (4) an infinite horizon and logarithmic utility. The two derivations of the CAPM in a multiperiod context, (3) and (4), are credited to Fama (1970b) and Rubinstein (1976), respectively. Moreover, Black (1972) derives a version of the CAPM that does not rely on the assumption of borrowing and lending at a risk-free rate but, instead, assumes unrestricted short sales of risky assets. This version of the CAPM differs from that in Eq. (2.6) only in that R_f is substituted by the expected return on assets that have a market beta equal to zero.

Early empirical tests of the CAPM focus on three hypotheses: (1) All assets show a linear relation between their expected returns and their market betas and no other variable provides any explanatory power over these returns, (2) the market beta risk price is positive, and (3) the expected returns on assets whose market betas are equal to zero are equal to the risk-free rate and the market beta risk price is the difference between the expected market return and the risk-free rate. These hypotheses are tested by either cross-sectional or time-series regressions and the third hypothesis does not apply to Black's (1972) version of the CAPM (e.g., Fama and French, 2004).

The first tests, such as those of Lintner (1965b), plot or regress the average returns of individual US stocks against their estimated betas. These early tests were, however, unsuccessful, since they involve (at least) three problems. First, betas of individual assets are unstable over time and measured with error, leading to bias when they are used as right-hand side variables in regressions. Second, individual stock returns are very volatile, so that one is unable to identify statistically significant differences in average returns (Cochrane, 2005, Sec. 20.2). Third, the residuals of such cross-sectional regressions are correlated, so that the standard errors of the slope coefficients are biased when they are estimated via the usual ordinary least squares methodology. The first two issues are addressed by Blume (1970), Friend and Blume (1970), Black et al. (1972), and Fama and MacBeth (1973), among others, who analyze portfolios of assets instead of individual securities. Returns on diversified portfolios are less volatile and portfolio betas are measured more accurately. Since the formation of portfolios, however, reduces the range of betas and, consequently, might reduce the statistical power of these tests, portfolios are often formed by sorting assets on their betas. The third issue is addressed by Fama and MacBeth (1973), who conduct month-by-month cross-sectional regressions of monthly returns on betas rather than estimate a single cross-sectional regression of average returns on betas. The authors use the time-series means of the monthly regression coefficients as well as the standard errors of these means to test the CAPM. Beyond that, Jensen (1968) proposes testing the CAPM by conducting asset-by-asset time-series regressions of excess returns on an intercept and the excess market return. If the CAPM is specified correctly, the intercept term, which is also known as Jensen's alpha, is zero for all assets (e.g., Fama and French, 2004).

The early cross-sectional tests of the CAPM of Douglas (1968), Black et al. (1972), Miller and Scholes (1972), Blume and Friend (1973), and Fama and MacBeth (1973) confirm hypotheses (1) and (2) but reject hypothesis (3). Although the relation between beta and average return seems to be positive (so that hypothesis (2) is confirmed), it is too flat, that is, the estimated slope coefficient is smaller than the average excess return on the market (which is typically proxied as the average return on a portfolio of US stocks in excess of the one-month US Treasury bill rate) and the intercept is greater than the risk-free interest rate (typically measured as the one-month US Treasury bill rate). Early time-series tests of the CAPM, such as those of Friend and Blume (1970), Black et al. (1972), and Stambaugh (1982), verify this conclusion in that the intercepts of low-beta assets tend to be positive, while the intercepts of high-beta assets tend to be negative. Moreover, Fama and MacBeth (1973) test whether additional variables, such as squared market betas (to test the linear relationship between expected returns and market beta) and residual variances from time-series regressions of returns on the market return (to examine whether market beta is sufficient to explain expected returns), have explanatory power over average returns beyond market beta. Their results suggest that these variables do not provide any explanatory power in addition to market beta, so that hypothesis (1) cannot be rejected (e.g., Fama and French, 2004).

Overall, the CAPM performs quite well in early empirical tests (Cochrane, 2005, Sec. 20.2). Although hypothesis (3) is consistently rejected, hypotheses (1) and (2) seem to hold. Hence, these early tests suggest that Black's (1972) version of the CAPM (which does not rely on the third hypothesis) is an especially good explanation of expected returns and "pushed the CAPM to the forefront of finance" (Fama and French, 2004, p. 35).

Roll (1977) criticizes the early tests of the CAPM, stating that a test of the model is in effect impossible, since the true market portfolio, which consists of every asset in the economy, is unobservable. Hence, tests of the CAPM will always have to use proxies such as stock market indices. The author emphasizes that Sharpe (1964) and Lintner (1965a) start their derivation of the CAPM with assumptions that imply that investors hold portfolios that are mean–variance efficient, in accordance with Markowitz (1952, 1959). Then, the CAPM adds assumptions that guarantee that the market portfolio is mean–variance efficient (and, hence, on the minimum variance frontier). The relationship between expected returns and beta in Eq. (2.6) is then just the application to the market portfolio of an algebraic relation that holds for any minimum–variance portfolio—and, consequently, for any mean–variance-efficient portfolio (e.g., Fama 1996; Fama and French, 2004). One conclusion of Roll's critique is that hypotheses (1) and (2) will be confirmed if the portfolio that is used as a proxy for the market portfolio is on the minimum–variance frontier that can be constructed from this proxy and the testing assets. According to the early CAPM tests mentioned above, the acceptance of these two hypotheses should confirm Black's (1972) version of the CAPM. This conclusion is, however, not possible, since the portfolio used to proxy for the market portfolio does not represent the true market portfolio. Hence, the bottom line of the early empirical tests of the CAPM, which confirm hypotheses (1) and (2), is that the market proxies they use seem to be on the minimum–variance frontier, rather than concluding that Black's (1972) version of the CAPM is the true asset pricing model (e.g., Fama and French, 2004).

2.2.1.2. The Intertemporal CAPM

Researchers in the early 1970s were not only engaged in empirically testing the CAPM but also highly active on the theoretical front, weakening the strong and rather unrealistic assumptions of the CAPM to develop more complex asset pricing models that more closely describe reality. Merton (1973) proposes an intertemporal, continuous-time model that is a natural extension of the CAPM and that is termed the Intertemporal CAPM (ICAPM). Long (1974) develops a discrete-time version of the ICAPM.

Since the CAPM is, in its original form, a two-period model, investors only care about the wealth they obtain through their portfolio in the second period, which they completely consume. In contrast, the ICAPM assumes investors who live more than two periods, trade continuously, and maximize the expected utility of their lifetime consumption. These investors do not have to completely consume the wealth their portfolio produces in the second (or next) period. Instead, they have the ability to reinvest some fraction of it. Consequently, they care not only about the wealth their portfolio produces in the next period, but also the opportunities they will have to consume or reinvest it. Such consumption investment opportunities are the relative prices of consumption goods (when the model assumes an economy with multiple goods) and the conditional distribution of asset returns. These are described by state variables. The bottom line of the ICAPM is that investors are concerned with both their wealth and state variables that capture their set of consumption investment opportunities. If consumption investment opportunities vary stochastically through time, expected returns depend not only on their sensitivity to variation in the market return, but also on their covariation with state variables (e.g., Cochrane, 2005, Sec. 9.2; Fama and French, 2004). If one assumes, instead, consumption investment opportunities that move deterministically through time, one is back to the CAPM of Eq. (2.6), as shown by Fama (1970b).

One discrete-time version of the ICAPM is derived by Fama (1996). The intuition for this derivation is analogous to the logic of the CAPM emphasized by Roll (1977). Fama states that ICAPM investors invest in mean–variance-efficient portfolios (as do Markowitz and CAPM investors who are only concerned with wealth uncertainty), as well as in portfolios that hedge uncertainty about shifts in consumption investment opportunities. Fama introduces the term *multifactor efficient* to characterize the ICAPM investors' overall portfolios. To be more specific, these investors minimize the variance of their overall portfolios, given a target expected portfolio return and a target vector of slope coefficients on the state variables (i.e., sensitivities with regard to variation in the state variables), since they are risk averse with respect to wealth uncertainty. The portfolios obtained that way are *multifactor minimum–variance*. Since ICAPM investors like wealth, they choose from this subset of portfolios those that maximize expected returns, given their return variance and their vector of slope coefficients on the state variables—and the resulting portfolios are multifactor efficient. Just as the assumption of market clearing prices in the CAPM implies that the market portfolio is mean–variance efficient, the assumption of market equilibrium in the ICAPM requires that the market portfolio be multifactor efficient. The risk–return relation of the ICAPM is then the application of an algebraic condition that holds in any multifactor minimum–variance portfolio (and, consequently, any multifactor-efficient portfolio) to the market portfolio. The ICAPM pricing equation is then a natural generalization of Eq. (2.6):

$$E(R_i) - R_f = \beta_{im}\left[E(R_m) - R_f\right] + \sum_{s=1}^{S} \beta_{is}\left[E(R_s) - R_f\right], \tag{2.8}$$

where R_s is the return on a portfolio that mimics uncertainty about state variable s and β_{im} and β_{is} are the slope coefficients from a multiple regression of R_i on R_m and R_s, respectively (Fama, 1996). Observe that this expression only differs from the pricing equation of the CAPM in Eq. (2.6) in the term $\sum_{s=1}^{S} \beta_{is}\left[E(R_s) - R_f\right]$. It shows that the pricing equation of the ICAPM considers two sources of risk: The first one is considered by the CAPM and is associated with changes in current aggregate wealth. The second one captures the risk of changes in the state variables, that is, the risk of shifts in consumption investment opportunities.

Another discrete-time version of the ICAPM is developed by Campbell (1993). This derivation is based on a dynamic consumption-based model with a representative investor who has Epstein–Zin (1989, 1991) recursive preferences. The resulting relation between expected return and risk is a log-linear approximation and is a multifactor model with the market return and shocks to variables that predict the market return (called discount rate news) as factors. The major drawback of the ICAPM is that it does not directly identify the state variables, which makes it, initially, a purely theoretical framework.

2.2.1.3. CAPM anomalies and the Fama–French three-factor model

Researchers in the 1970s were not content with the results of the early CAPM tests and the success of Black's (1972) version of the CAPM, so they kept testing the model. Much evidence that starts emerging in the late 1970s proves them right by suggesting that there is much variation in expected returns on US stocks that cannot be captured by market beta. These deviations from the CAPM are termed anomalies.

Basu (1977, 1983) finds that, controlling for market beta, expected stock returns rise with earnings–price ratios. Moreover, Banz (1981) provides evidence for a size effect: Small stocks, that is, stocks with low market capitalization, have higher average returns than predicted by the CAPM. Furthermore, Bhandari (1988) documents that leverage, defined as the ratio of the book value of debt to the market value of equity, is positively related to average returns, controlling for market beta. Additionally, Stattman (1980) and Rosenberg et al. (1985) observe a positive relationship between average returns and the book-to-market ratio that cannot be explained by market beta, while DeBondt and Thaler (1985) discover a reversal in long-term returns, that is, stocks with low long-term past returns tend to have higher returns in the future. Jegadeesh (1990) finds that a stock's expected return is negatively related to its lagged return. Beyond that, Reinganum (1981), Stambaugh (1982), and Lakonishok and Shapiro (1986) observe that the relation between market beta and average returns is actually flatter than in the early tests of the CAPM and even disappears in a more recent sample period.

The list of further characteristics that seem to impact expected returns in a way that is unrelated to market beta is long: La Porta (1996) finds a negative relation between expected returns and analysts' earnings growth forecasts, while Sloan (1996) provides evidence of a negative relation between expected returns and levels of accounting accruals. Frankel and Lee (1998) show that expected returns are positively related to the ratio of a stock's fundamental value estimated from Institutional Brokers' Estimate System consensus forecasts and a residual income model to its price. More recently, Titman et al. (2004) find that a firm's expected return is negatively related to its capital investments, while Zhang (2006) relates average returns to information uncertainty that results from, for instance, low analyst following. Moreover, Hou and Robinson (2006) find that firms in more concentrated industries show lower average returns and Daniel and Titman (2006) argue that expected returns are negatively related to a stock's intangible return, defined as the component of its past return that is unrelated (orthogonal) to the firm's past performance. The authors state that both the book-to-market ratio and equity issuance are related to intangible returns. Furthermore, Cooper et al. (2008) find that a firm's expected return is negatively related to its annual book asset growth rate, while Fang and Peress (2009) provide evidence that stocks with no media coverage exhibit higher returns than stocks with high media coverage. Finally, Fu (2009) finds a significantly positive relation between conditional idiosyncratic volatilities and expected returns.

More CAPM anomalies found over the years involve small trade order flows (Hvidkjaer, 2008; Barber et al., 2009); bankruptcy risk (Dichev, 1998; Campbell et al., 2008); economic links between firms (Cohen and Frazzini, 2008); shareholder rights (Gompers et al., 2003);[5] measures of private information, such as the probability of information-based trading (Easley et al., 2002) and the geographical distance of investors (Coval and Moskowitz, 2001); the delay with which a stock's price responds to information (Hou and Moskowitz, 2005); short-sale constraints (Jones and Lamont, 2002; Asquith et al., 2005); profitability, measured by the ratio of gross profits to assets (Novy-Marx, 2013); and investment (Aharoni et al., 2013).[6] Finally, momentum, several forms of illiquidity, as well as sensitivities to innovations in state variables inspired by the ICAPM are found to explain expected returns beyond market beta. I devote separate sections to these characteristics later in this review.

Fama and French (1992) update and synthesize much of the evidence on CAPM anomalies before 1992 regarding NYSE, AMEX, and NASDAQ stocks over the sample period 1963 to 1990. Among other things, they observe that different price ratios, such as book-to-market and earnings-to-price ratios, give similar signals with regard to expected returns. Stocks with high ratios of a fundamental such as book value to price are termed value stocks and stocks with low ratios of fundamentals to price are termed growth stocks. The difference in average returns on value and growth stocks is known as the value effect. Fama and French (2004, p. 36) emphasize that their 1992 paper is "marking the point when it is generally acknowledged that the CAPM has potentially fatal problems."

Many of these characteristics can be linked to the factors presented in Sec. 2.1.1 that seem to predict stock returns in the time series. In particular, many stock characteristics that seem to describe the cross-sectional variation in average returns also seem to predict returns in the time series. Examples are the earnings-to-price ratio, the book-to-market ratio, lagged returns, capital investments or the ratio of investment to capital, volatility, and momentum. Cochrane (2005, Sec. 20.2) notes that the value effect is the cross-sectional analogy to the predictability of returns in the time series by the dividend–price ratio and other price ratios. This is also emphasized in Cochrane (2011, p. 1062):

> Is value a "time-series" strategy that moves in and out of a stock as that stock's book-to-market ratio changes, or is it a "cross-sectional" strategy that moves from one stock to another following book-to-market signals? Well, both, obviously. They are the same thing.

Because of these similarities, Cochrane (2011) suggests that an asset's expected return should be generally regarded as a function of its characteristics. These characteristics can change both through time and across assets, so that researchers are actually examining big panel data-forecasting regressions, instead of estimating expected returns separately in the time series and cross section.

Fama and French (1993) present a three-factor asset pricing model (henceforth FF3) that seeks to explain the patterns in average returns associated with size and value. FF3 is an empirical asset pricing model. Classic asset pricing models such as the CAPM and ICAPM work forward, that is, from theoretical assumptions to predictions about how one should

[5] The findings of Gompers et al. (2003) are refuted to some extent by Johnson et al. (2009).

[6] For a more detailed summary of much of these anomalous effects, see, for instance, Subrahmanyam (2010).

measure risk and to the relation between expected return and risk. In contrast, empirical asset pricing models work backward. They take observed patterns in average returns as given and suggest models to explain them. FF3 takes the size and value effect as given and seeks to capture these two effects (Fama and French, 2013). To be specific, the model's prediction about the relation between expected return and risk is

$$E(R_i) - R_f = b_i\left[E(R_m) - R_f\right] + s_i E(SMB) + h_i E(HML), \tag{2.9}$$

where SMB (small minus big) is the difference in the returns on a diversified portfolio of small stocks and a diversified portfolio of big stocks and HML (high minus low) is the difference between the returns on diversified portfolios of stocks with high and low book-to-market ratios. Moreover, b_i, s_i, and h_i are the slope coefficients in the following time-series regression:

$$R_{i,t} - R_{f,t} = a_i + b_i\left(R_{m,t} - R_{f,t}\right) + s_i SMB_t + h_i HML_t + e_{i,t}, \tag{2.10}$$

where $R_{i,t}$ is the return on asset i in month t, $R_{f,t}$ is the risk-free return in month t (with similar notation for the remaining returns), a_i is a constant, and $e_{i,t}$ is a residual term. Fama and French (1993) evaluate their model using the time-series regression test of Black et al. (1972), that is, they test whether a_i in Eq. (2.10) is zero for all assets i. Their data sample consists of NYSE, AMEX, and NASDAQ stocks over the period 1963 to 1991.

Fama and French (1993) construct SMB and HML as follows: In June of each year t they sort all stocks into two size groups (small and big), where the median NYSE market capitalization is used as the breakpoint. Additionally, they sort all stocks into three groups based on their book-to-market ratios (low, medium, and high), using the 30th and 70th percentiles of the book-to-market ratios of NYSE stocks as breakpoints. To ensure the availability of accounting data to investors, they define the book-to-market ratio in t as the ratio of book common equity for the fiscal year ending in calendar year $t-1$ to market equity at the end of December $t-1$. Then, Fama and French construct six portfolios from the intersections of the two size groups and the three book-to-market groups (small–low, small–medium, small–high, big–low, big–medium, and big–high). Afterward, they calculate monthly value-weighted returns on each of these six portfolios from July of year t to June of $t+1$. The portfolios are reformed in June of $t+1$. The return on SMB is then computed as the simple average of the returns on the three small-stock portfolios minus the simple average of the returns on the three big-stock portfolios. Similarly, the return on HML is calculated as the difference between the simple average of the returns on the two high–book-to-market portfolios and the simple average of the returns on the two low–book-to-market portfolios. Finally, the market portfolio is proxied by the value-weighted portfolio of the stocks that are included in the six portfolios to construct SMB and HML, plus stocks with negative book equity that were excluded from these portfolios. The risk-free rate is measured by the one-month US Treasury bill rate.

Fama and French (1993) propose a set of 25 portfolios of stocks sorted by size and book-to-market ratio as testing assets (SBM25). These portfolios produce a wide range of average returns that represent the size and value effects. Fama and French construct them similarly to the six portfolios used to construct SMB and HML, employing NYSE, AMEX, and NASDAQ stocks. Even today, these portfolios are one of the most challenging portfolios in asset pricing. For this reason, they have become the benchmark test for competing asset pricing models (e.g., Petkova, 2006). As a robustness check, Fama and French use portfolios formed on the earnings–price and dividend–price ratios. Moreover, they use seven bond portfolios to see whether FF3 is also able to explain the variation in average returns across government and corporate bonds with different times to maturity and ratings.

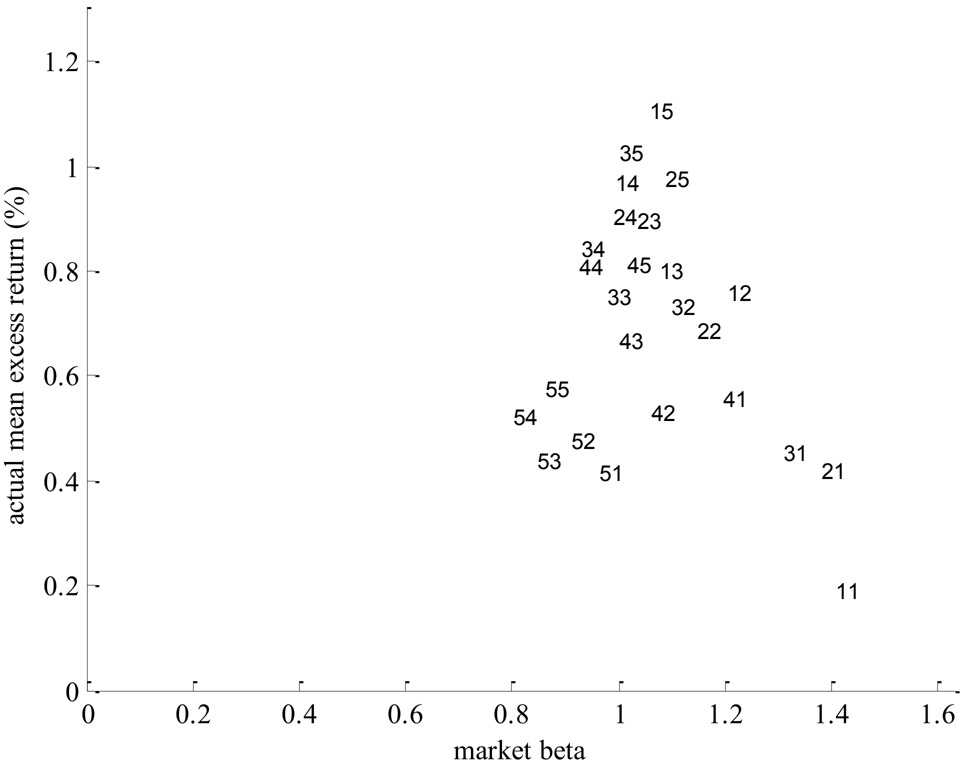

Figure 1. Mean excess returns versus market betas (SBM25).
This figure plots the market betas of 25 portfolios of NYSE, AMEX, and NASDAQ stocks sorted by size and book-to-market ratio against their percentage monthly mean returns in excess of the one-month US Treasury bill rate, measured over the sample period July 1963 to December 2012. In this figure, 11 denotes the portfolio that consists of the smallest stocks with the lowest book-to-market ratios (small–growth stocks), 15 denotes the portfolio of the smallest stocks with the highest book-to-market ratios (small–value stocks), and 55 denotes the portfolio of the biggest stocks with the highest book-to-market ratios (big–value stocks). The notation of the remaining portfolios follows this scheme. The data are obtained from Kenneth R. French's website,[7] which also provides a detailed description of the construction of these portfolios.

Figure 1 plots the average excess returns on SBM25 against their market betas and provides an updated version (with data from 1963 to 2012) of Cochrane's (2005) Figure 20.9. At first glance, one can see that there is a lot of variation in average excess returns that is unrelated to market beta. If one takes a closer look at the figure, one even observes that variation in the book-to-market ratio (within a given size group) goes along with variation in average excess returns that is *negatively* related to market beta. For this reason, the value effect makes the

[7] See http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

CAPM “a disaster when confronted with these portfolios” (Cochrane, 2005, Sec. 20.2, p. 439).

Fama and French (1993) observe that FF3 performs considerably better than the CAPM in explaining the average excess returns on SBM25. Only three of the 25 intercepts a_i from Eq. (2.10) differ from zero by more than 0.2% per month and 16 are within 0.1% and zero. Relatively large intercepts (in absolute terms) are shown by the portfolios of stocks in the lowest book-to-market quintile (growth stocks). Among these, the portfolio with the smallest stocks shows—with an intercept of -0.34% per month—an average return that is statistically significantly *too low* to be explained by FF3 (the t-statistic of the intercept is -3.16), while the biggest stocks show average returns that are *too high* to be captured by the model (the intercept is 0.21% per month, with a t-statistic of 3.27). Fama and French use the F-test proposed by Gibbons, Ross, and Shanken (1989), or GRS—that is, the GRS test and the GRS statistic—to evaluate the null hypothesis that the intercepts on the 25 stock portfolios and the seven bond portfolios are jointly equal to zero. The authors observe that this null hypothesis is rejected by the GRS test. Hence, FF3 is already in its original paper rejected by the data. Nevertheless, FF3 does quite a good job (relative to the CAPM) explaining the size and value effects and has become one of the most popular asset pricing models, dominating empirical research for many years (Cochrane, 2005, Sec. 20.2).

Figure 2 plots the actual average excess returns on SBM25 (see Figure 1) against their expected excess returns as predicted by FF3. The figure is similar to Figures 20.12 and 20.13 of Cochrane (2005), but uses data from 1963 to 2012. If FF3 perfectly described the actual mean excess returns, all portfolios would lie on a 45° line. One can see that the model fits quite well. The main exceptions are the small–growth portfolio (denoted by 11) and the big–growth portfolio (denoted by 51), as observed by Fama and French (1993). Hence, even 20 years after Fama and French’s study, the model seems to still be rejected mainly because of these two portfolios.

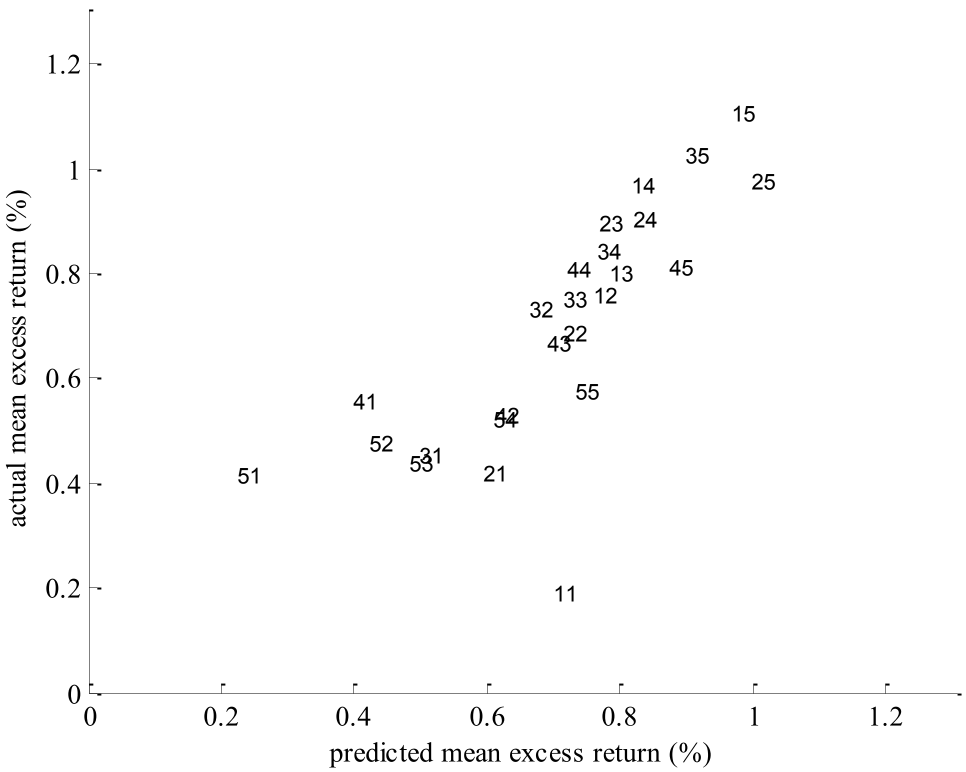

Figure 2. Actual mean excess returns on SBM25 versus mean excess returns predicted by FF3.
The figure plots the percentage monthly mean excess returns as predicted by FF3 of 25 portfolios of NYSE, AMEX, and NASDAQ stocks sorted by size and book-to-market ratio against their actual percentage monthly mean returns in excess of the one-month US Treasury bill rate, measured over the sample period July 1963 to December 2012. In this figure, 11 denotes the portfolio that consists of the smallest stocks with the lowest book-to-market ratios (small–growth stocks), 15 denotes the portfolio of the smallest stocks with the highest book-to-market ratios (small–value stocks), and 55 denotes the portfolio of the biggest stocks with the highest book-to-market ratios (big–value stocks). The notation of the remaining portfolios follows this scheme. The data are obtained from Kenneth R. French's website, which also provides a detailed description of the construction of these portfolios and of the three explanatory factors.

Fama and French (1993) also propose a five-factor model (FF5) by adding two bond market factors to FF3. The factors are the long-term return spread, TMR, and the default return spread, DFR, resulting in the following time-series regression to test:

$$\begin{aligned} R_{i,t} - R_{f,t} = a_i + b_i\left(R_{m,t} - R_{f,t}\right) + s_i SMB_t + h_i HML_t \\ + m_i TMR_t + d_i DFR_t + e_{i,t}, \end{aligned} \tag{2.11}$$

where TMR is proxied by the difference between the monthly return on a long-term US government bond and the US Treasury bill rate (which is measured in the previous month) and the proxy for DFR is the difference between the return on a portfolio of long-term US corporate bonds and the long-term US government bond return. Fama and French observe, however, that the model produces a GRS statistic that is worse than that obtained with FF3.

Fama and French (1996) do further tests on FF3 and conclude that many of the CAPM anomalies discovered so far are captured by their model. To be specific, they observe that the model captures the variation in average returns of portfolios sorted by earnings–price ratios, cash flow–price ratios, as well as sales growth. Moreover, they show that FF3 explains the reversal in long-term returns found by DeBondt and Thaler (1985). However, they conclude that FF3 cannot capture the momentum effect documented by Jegadeesh and Titman (1993) (more on this in Sec. 2.2.1.4).

How can CAPM anomalies, especially size and value, and the empirical success of FF3 be interpreted? Fama and French (1996) discuss three economic interpretations—risk based, behavioral, and spurious—as well as one purely algebraic explanation.

The **risk-based interpretation** assumes efficient markets with rational investors who believe that small stocks are riskier than big stocks, that value stocks are riskier than growth stocks, and that these risks are not captured by a stock's market beta. This interpretation is supported by, for instance, Chan and Chen (1991) and Perez-Quiros and Timmermann (2000), who show that the high average returns on small stocks are related to their relatively high exposure to variations in credit market conditions, and Vassalou and Xing (2004), who state that both the size and value effects are connected to systematic default risk. Similarly, Zhang (2005) shows that the value effect appears naturally in a neoclassical industry equilibrium model with aggregate uncertainty and rational expectations.

A risk-based explanation for the success of FF3 is that investors rationally price assets according to a three-factor ICAPM, or APT, but not according to the CAPM. Indeed, Cochrane (2005, Sec. 20.2) states that "in retrospect, it is surprising that the CAPM worked so well for so long," since, at least since Merton's (1973) ICAPM, one knows the theoretical possibility that explaining the cross section of average returns requires more than market beta. For this explanation to work, there must be some non-diversifiable risks that are proxied by HML and SMB and that make market participants worry about investing in stocks that do poorly when HML and SMB do poorly although the market portfolio does well (Cochrane, 2005, Sec. 20.2). The results of Fama and French (1993) show that there is covariation in returns that is related to size and the book-to-market ratio and that cannot be explained by the assets' sensitivities to the market return. In particular, both value stocks and small stocks seem to move together, since both kinds of stocks have high betas on HML and SMB, respectively. This comovement among small and value stocks supports both the ICAPM and APT stories.

To interpret FF3 as an ICAPM, one has to identify the state variables behind SMB and HML with which investors are concerned. Fama and French (1993, 1996) argue that $E(HML)$ might be the premium for a state variable that is related to relative distress. They base their interpretation on the finding of Fama and French (1995) that high book-to-market equity is associated with firms with persistently low earnings (and vice versa), that is, with firms that are in or near financial distress, and on Fama and French's (1997, hereafter FF97) discovery that industries' loadings on HML vary through time in accordance with periods of industry strength or distress. However, an *individual* firm or industry's financial distress cannot be a *systematic* risk factor that requires a higher expected return, since this risk can be diversified away in a portfolio (Cochrane, 2005, Sec. 20.2). Instead, there must be a state variable associated with relative distress that is consistent with the ICAPM—that is, an economy-wide variable that investors systematically seek to hedge—and these hedging demands must not balance out among individual investors, so that they affect the *average* investor. Fama and French (1996) suggest a relative-distress premium that is based on investors who hedge their human capital. In particular, the authors argue that distressed firms might do *very* badly, that is, more badly than non-distressed firms, in the event of a negative shock to the firm's prospects. Consequently, workers of distressed firms with specialized human capital will more likely face a negative shock to the value of their human capital than workers of non-distressed firms, that is, they will more likely lose their jobs. To hedge their human capital, workers of distressed firms will avoid holding (or even bet against) stocks of their own firms. If negative shocks to prospects are correlated across (distressed) firms, that is, if they are likely to happen economy-wide at the same time, all workers of distressed firms will probably avoid the stocks of all distressed firms. Such systematic investor behavior may result in a state-variable risk premium on the expected returns on distressed stocks, that is, on stocks with high book-to-market ratios. However, Fama and French (1996) admit that their story is disputable. For instance, it seems that HML is "unfortunately" not strongly correlated with other measures of economy-wide financial distress, as pointed out by Cochrane (2005, Sec. 20.2).

An interpretation of FF3 as an APT seems to be more conclusive. According to the APT, FF3 would have to hold if the R^2 in the time-series regressions of the testing assets' excess returns on the three Fama–French factors are all 100%, that is, if one can perfectly replicate the returns by the three factors incorporated in FF3. Otherwise, there would be arbitrage opportunities. Cochrane (2005, Sec. 20.2) emphasizes that the R^2 estimates in the time-series tests with SBM25 are almost all above 90%. Consequently, a large part of the time-series variation in the ex post returns on these 25 portfolios can be explained by the three factors included in FF3, which indicates that FF3 might be a good APT (Cochrane, 2005, Sec. 20.2).

Several other studies support a rational explanation of FF3's empirical success. For instance, Liew and Vassalou (2000) and Vassalou (2003) provide evidence that SMB and HML mimic news related to future gross domestic product growth, which suggests that these factors reflect certain macroeconomic risks. Moreover, Lettau and Ludvigson (2001b) and Vassalou (2003) show that the consideration of macroeconomic risk reduces the explanatory power of SMB and HML. Furthermore, Petkova (2006) shows that both SMB and HML lose their information content for the cross section of SBM25 average returns when these two factors are included in a model with several predictors of future investment opportunities, which supports the ICAPM interpretation of FF3. The author concludes that SMB and HML proxy for innovations in state variables that forecast investment opportunities. A similar conclusion is drawn by Hahn and Lee (2006). In addition, MSC find that FF3 is (to some extent) consistent with the ICAPM within the US stock market. The latter three studies are presented in more detail in Secs. 2.2.1.6 and 2.2.1.7. More evidence that supports a risk-based interpretation of FF3 is provided by Lewellen (1999), among others.

The **behavioral interpretation** argues that the size and value anomalies as well as the empirical success of FF3 are the results of inefficient markets and investor irrationality. In particular, researchers such as DeBondt and Thaler (1987) argue that investors might behave irrationally and simply overreact to new information. For instance, the prices of value stocks might just be *too* low and the prices of growth stocks might be *too* high relative to their fundamentals. A correction of such an overreaction of investors would lead to high returns on value stocks and low returns on growth stocks. With regard to FF3, $E(HML)$ might consequently be too large to be caused by rational pricing and the HML portfolio might be close to an arbitrage opportunity, as discussed by Lakonishok et al. (1994), Haugen (1995), MacKinlay (1995), and others. Additionally, Lakonishok et al. (1994) argue that the positive value of $E(HML)$ might arise because investors simply dislike distressed stocks and underprice them rather than rationally and systematically hedge a state variable related to relative distress. Beyond that, Daniel and Titman (1997) provide evidence that it is the stocks' specific characteristics that explain their expected returns, rather than their covariances with common factors. The authors conclude that there are no common risk factors associated with size and value and none of the FF3 factors actually commands a risk premium. Hence, their evidence suggests that the size and value premiums cannot be viewed as compensation for state variable or factor risk.

The **spurious interpretation** of the size and value anomalies and FF3's empirical success suggests that the CAPM in effect holds, but that the model is spuriously rejected in empirical tests. Reasons for rejection of the CAPM might be a survivorship bias among the firms with which the model is tested (Kothari et al., 1995), poor proxies for the market portfolio (Roll, 1977), or other statistical problems. For instance, Berk et al. (1999) and Gomes et al. (2003) suggest that the results of Fama and French (1992) might simply be due to problems in the measurement of market beta. Moreover, the CAPM anomalies might be the result of data snooping activities, that is, researchers might be extensively searching through the same data set (the CRSP database) for patterns that are inconsistent with the CAPM. The resulting anomalies may just be the sample-specific results of random chance, or "luck." This explanation is suggested by Lo and MacKinlay (1990), Black (1993), and MacKinlay (1995). Such an interpretation is supported by the fact that many anomalies seem to decline over time. For instance, the value effect was cut roughly in half in the US in the 1990s and the size effect seems to even vanish completely after 1980, when Banz (1981) discovered the effect and the first small-stock funds emerged that made trading and exploitation of the effect possible for the average investor. Nevertheless, the value effect is still apparent and the whole story might be told with value alone (Cochrane, 2005, Sec. 20.2; Fama and French, 2012). Finally, Ferson et al. (1999) argue that characteristic-sorted portfolios may behave like risk factors even when the characteristics are not related to risk. However, Cochrane (2005, Sec. 20.2) and others defend FF3 in this regard: Although the model explains portfolios sorted by two characteristics (size and book to market) using factors that are portfolios constructed on the basis of the same two characteristics (although with a rougher grid), FF3 is not a tautology.

The **purely algebraic (minimalistic) interpretation** of FF3's empirical success builds on Roll's (1997) critique of CAPM tests and on Fama's (1996) derivation of the ICAPM. In particular, one can argue that the right-hand side portfolios of Eq. (2.9) are in the set of three-factor minimum–variance portfolios that can be formed from these right-hand side portfolios and the testing assets (i.e., they span these portfolios) and the linear relation between the testing assets' returns and the right-hand side portfolios of FF3 holds simply for this reason (Fama and French, 1996).

2.2.1.4. Momentum and the Carhart four-factor model

Jegadeesh and Titman (1993) find that portfolios that buy short-term winners and sell short-term losers produce significant average returns that are not due to risk measured by market beta. This effect is termed momentum. Momentum is similar to the long-term reversal anomaly pointed out by DeBondt and Thaler (1985), but the sign is the opposite: Long-term losers tend to do well in the long term, while short-term losers tend to continue to do badly in the short term.

In particular, Jegadeesh and Titman (1993) examine several trading strategies that are based on past returns, using data for NYSE and AMEX stocks over the sample period 1965 to 1989. In each month t, they sort stocks into 10 equal-weighted portfolios on the basis of their returns over the past J months, where the portfolio that consists of the stocks with the lowest past returns is called the losers portfolio and the portfolio that includes the stocks with the highest past returns is the winners portfolio. In each month t, their strategy buys winners and sells losers and this trading position is held for the next K months. Since this is done each month, their overall strategy consists of several trading positions with overlapping holding periods. Jegadeesh and Titman examine strategies that use values of three, six, nine, and 12 months for J and K (J and K do not have to be equal). They find that most of these strategies yield average returns that are significantly positive. Moreover, they document that these returns cannot be explained by market beta. For instance, the strategy for $J = K = 6$ produces an average excess return of 12.01% per year but a slightly negative beta on the market portfolio.

Figure 3 plots the average excess returns on 25 portfolios of NYSE, AMEX, and NASDAQ stocks sorted by size and momentum (SM25) against their market betas that are constructed similarly to SBM25 of Sec. 2.2.1.3. One can see that there is much variation in average excess returns that is unrelated or even negatively related to market beta. For instance, the spread in average excess returns of the small–winner portfolio (denoted by 15) and the small–loser portfolio (denoted by 11) is around 1.4% per month, although the latter portfolio shows a higher market beta.

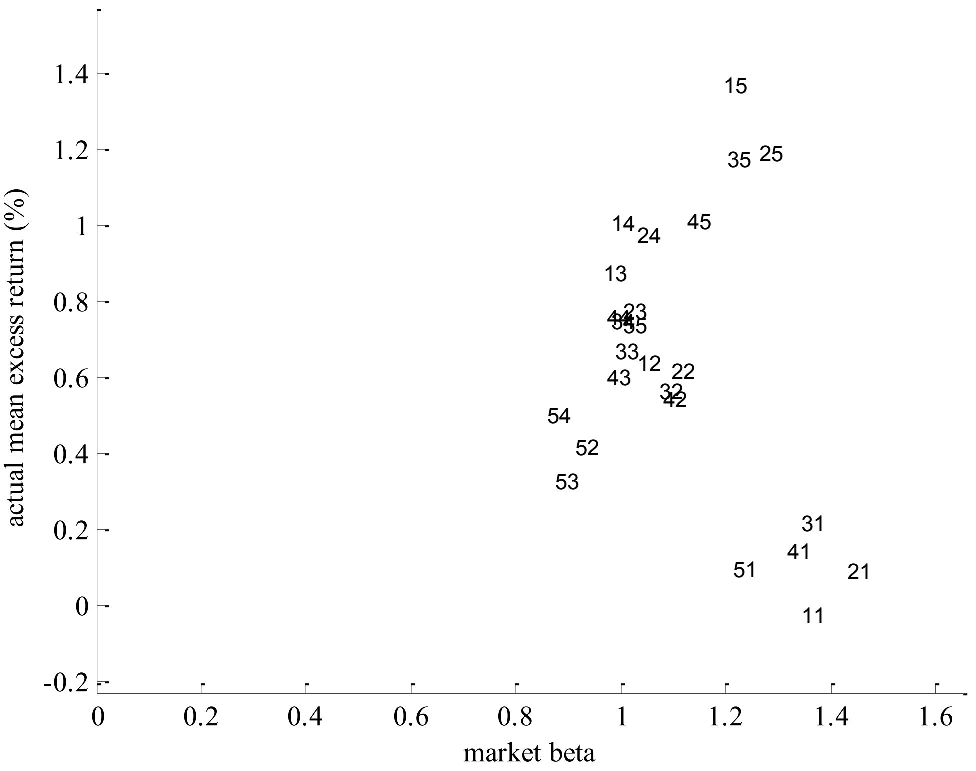

Figure 3. Mean excess returns versus market betas (SM25).
This figure plots the market betas of 25 portfolios of NYSE, AMEX, and NASDAQ stocks sorted by size and momentum against their percentage monthly mean returns in excess of the one-month US Treasury bill rate, measured over the sample period July 1963 to December 2012. In this figure, 11 denotes the portfolio that consists of the smallest stocks with the lowest returns over the past year (small–loser stocks), 15 denotes the portfolio of the smallest stocks with the highest returns over the past year (small–winner stocks), and 55 denotes the portfolio of the biggest stocks with the highest returns over the past year (big–winner stocks). The notation of the remaining portfolios follows this scheme. The data are obtained from Kenneth R. French's website, which also provides a detailed description of the construction of these portfolios.

How can momentum be interpreted? Cochrane (2005, Sec. 20.2) argues that momentum simply arises from the weak time-series predictability of monthly individual stock returns. Since such predictability is already observed in studies prior to that of Jegadeesh and Titman (1993) (e.g., Fama, 1965), Cochrane (2005, p. 447) states that "momentum is really a new way of looking at an old phenomenon." The buying and selling of portfolios of extreme winners and losers is what produces expected portfolio returns that are economically significant from tiny (but statistically significant) autocorrelations in individual stock returns. However, the author states that it is disputable whether momentum is exploitable after transactions costs, since momentum strategies, such as those proposed by Jegadeesh and Titman, require frequent trading. For example, Carhart (1997) concludes that momentum is not exploitable after transactions costs. While FF3 is able to capture long-term reversal, it is unable to explain momentum (Fama and French, 1996). In effect, past short-term losers tend to have rather low prices and seem to move with value stocks. Consequently, FF3 predicts that these stocks should have high average returns, but they behave in the opposite manner (Cochrane, 2005, Sec. 20.2).

Moskowitz and Grinblatt (1999) identify industries as a key source of momentum profits. They state that their finding may be supportive for a behavioral explanation of momentum. In particular, they argue that if momentum profits are driven behaviorally, for example, by investors that underreact to information, then these momentum trading strategies must be exposed to at least some factor risk that one cannot eliminate (e.g., some market beta). Otherwise, these strategies would represent an arbitrage opportunity to those investors who act rationally and would be exploited by them. The fact that industry momentum seems to generate much of individual stock momentum and that stocks within an industry move together makes momentum strategies not very well diversified. This, in turn, makes these strategies very different from an arbitrage and may support a behavioral explanation of momentum. Similarly, Hong et al. (2000) find that momentum profits decrease with size and analyst coverage and that this negative relationship between analyst coverage and momentum profits is greater for past loser stocks than for past winners. These findings support the hypothesis of Hong and Stein (1999), that momentum results from a gradual diffusion of firm-specific information across the investing public so that it is due to market inefficiencies.

Chordia and Shivakumar (2002) discover that portfolios sorted on momentum vary systematically in their sensitivity to a set of lagged variables that are known to predict stock returns in the time series, that is, to the dividend yield, the default yield spread, the Treasury bill rate, and the term structure spread (see Sec. 2.1.1). They show that momentum profits decrease significantly after controlling for these differences. The authors conclude that momentum payoffs can therefore be explained by time-series predictability and time-varying expected returns, that is, these payoffs are generated by buying portfolios with high conditional expected returns and selling portfolios that exhibit low conditional expected returns. Cooper et al. (2004) document that momentum profits depend on the state of the market, that is, they are much larger when they follow positive market returns than after negative ones. Specifically, the average monthly momentum profit following positive returns on the market is 0.93%, while the average profit after negative market returns is -0.37% for the sample period 1929 to 1995. Assuming that investor overconfidence is higher after a rising market, this finding speaks in favor of an overreaction theory of momentum.

Grinblatt and Moskowitz (2004) find that momentum returns depend on the degree of consistency of past returns, that is, momentum profits tend to be higher when the high past return of winner stocks is produced by a series of steady positive monthly returns, in contrast to a small set of extraordinary good months. Among other things, the authors argue that such consistency may be a measure for the inverse of volatility and therefore a proxy for risk that might affect expected returns. Furthermore, Hvidkjaer (2006) provides evidence that momentum is to some extent caused by the behavior of small traders. The author uses transaction data to measure the pressures of buying and selling loser and winner stocks and provides evidence that indicates that momentum results from both initial underreaction and delayed reaction among small traders but not among large traders. For instance, it seems that small traders keep buying loser stocks for up to one year before they start to sell them. Avramov et al. (2007) find that momentum profits are generated largely by firms with a low credit rating. In particular, the profitability of momentum strategies seems to be high among firms with low credit ratings, but virtually zero among the stocks of high-grade firms.

Moreover, the low-grade firms that drive the momentum profits discovered in the literature seem to represent less than 4% of the overall market capitalization of rated firms.

Overall, it seems that there are a variety of different explanations of momentum proposed in the literature, which disagree in parts, and momentum seems to be much more puzzling than the size and value anomalies presented earlier. Some of these interpretations are risk based, but the majority of them seem to be behavioral. In a recent interview, Fama even states that "of all the potential embarrassments to market efficiency, momentum is the primary one" (Fama and Litterman, 2012, p. 18).

Nevertheless, momentum stocks tend to move together. Consequently, returns on momentum portfolios can be explained by a common momentum factor, just as FF3 explains the returns on small and value stocks. Such a model is proposed by Carhart (1997), who adds a momentum factor, UMD, to FF3, leading to the following time-series regression (henceforth C):

$$R_{i,t} - R_{f,t} = a_i + b_i(R_{m,t} - R_{f,t}) + s_i SMB_t + h_i HML_t + u_i UMD_t + e_{i,t}, \qquad (2.12)$$

where UMD_t is the return on a diversified portfolio of the past year's winners minus the return on a diversified portfolio of the past year's losers, constructed similarly to SMB and HML. Figure 4 plots the actual mean excess returns from Figure 3 against their expected excess returns as predicted by C. All portfolios should lie on a 45° line for C to hold perfectly. The figure suggests that C does a quite good job. The greatest challenge seems to be the small–loser portfolio (denoted by 11), which shows an average return that is too low to be explainable by C.

Carhart (1997) originally suggested C as a tool to measure the performance of US mutual funds, so that UMD is in effect a pure performance attribution factor. Indeed, Carhart constructs UMD using the portfolios of mutual funds, not stock portfolios. In particular, C is proposed to be used to evaluate whether a fund manager does well because of stock picking abilities or because the fund manager is simply following a mechanical momentum strategy (Cochrane, 2005, Sec. 20.2). Nevertheless, many studies use C to explain the average returns on stock portfolios sorted by momentum. These studies construct UMD using stock portfolios. Because of this difference to Carhart's performance attribution factor, C is sometimes called the Fama–French four-factor model (e.g., Asness et al., 2013).

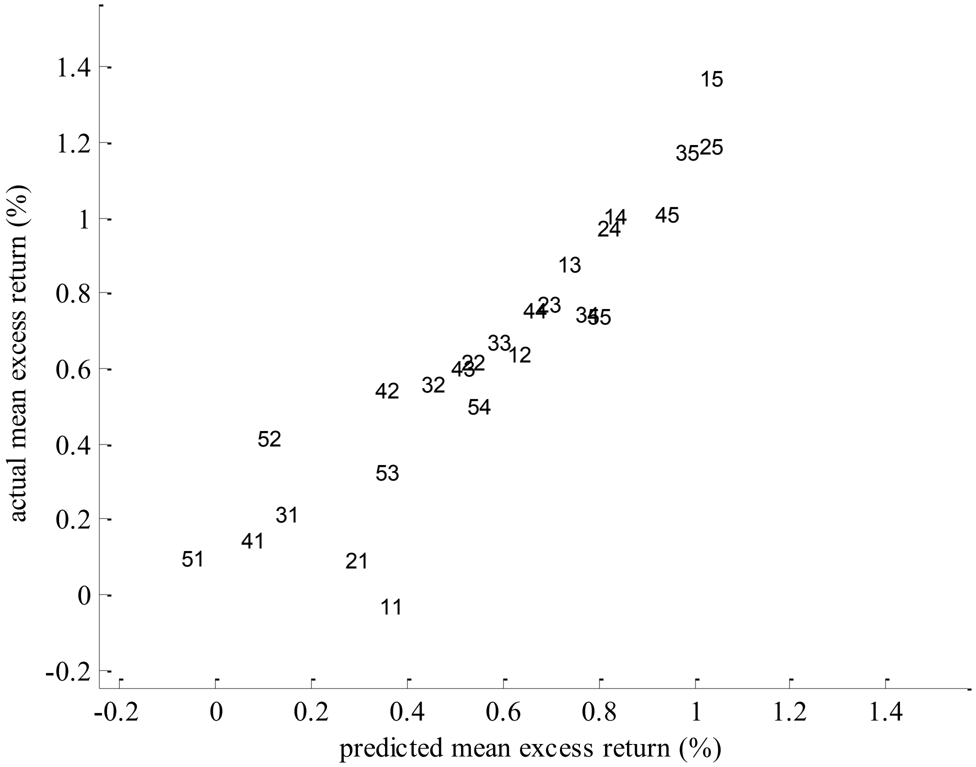

Figure 4. Actual mean excess returns on SM25 versus mean excess returns predicted by C.

The figure plots the percentage monthly mean excess returns as predicted by C for 25 portfolios of NYSE, AMEX, and NASDAQ stocks sorted by size and momentum against their actual percentage monthly mean returns in excess of the one-month US Treasury bill rate, measured over the sample period July 1963 to December 2012. In this figure, 11 denotes the portfolio that consists of the smallest stocks with the lowest returns over the past year (small–loser stocks), 15 denotes the portfolio of the smallest stocks with the highest returns over the past year (small–winner stocks), and 55 denotes the portfolio of the biggest stocks with the highest returns over the past year (big–winner stocks). The notation of the remaining portfolios follows this scheme. The data are obtained from Kenneth R. French's website, which also provides a detailed description of the construction of these portfolios and of the four explanatory factors.

However, since no underlying rational economic foundations of momentum are uncovered and since momentum seems to be much more elusive than size and value, adding the momentum factor to FF3 as a risk factor is subject to much criticism. For instance, Cochrane (2005, Sec. 20.2, p. 447) states that the use of UMD to explain momentum sorted portfolios "is so obviously ad hoc […] that nobody wants to add it as a risk factor."

2.2.1.5. Liquidity (risk)

A variety of studies examine the impact of liquidity (or illiquidity) on expected stock returns. One comprehensive review of this literature is given by Amihud et al. (2005). At first glance, liquidity seems to be quite an elusive concept. Nevertheless, a useful definition of asset liquidity, which many studies build on, is provided by Keynes (1930, p. 67): "An asset is more liquid than another if it is more certainly realisable at short notice without a loss." On this basis, liquidity is measured in several ways, for instance, using bid–ask spreads (Amihud and Mendelson, 1986), turnover (Datar et al., 1998), and trading volume (Amihud, 2002; Pástor and Stambaugh, 2003). More measures are provided by Lesmond et al. (1999), Liu (2006), and others. The literature on the relation between expected returns and liquidity can be roughly divided into two categories: The first category examines the impact of an asset's idiosyncratic (level of) liquidity on its expected return, while the second category analyzes aggregate liquidity as a systematic risk factor that affects expected returns (Pástor and Stambaugh, 2003). Empirical studies of the first category, such as those of Amihud and Mendelson (1986), Brennan and Subrahmanyam (1996), Brennan et al. (1998), and Datar et al. (1998), use a variety of measures for liquidity and typically show that illiquid stocks have higher average returns than liquid stocks. Hence, investors seem to require compensation for illiquidity. Analogously, Chordia et al. (2001) examine the impact of both the level and he variation of a stock's liquidity (i.e., the first and second moments of liquidity). They document (surprisingly) that average returns are negatively related to the second moment of liquidity. Studies that examine the relation between an asset's liquidity and its expected return or price (also) from a theoretical perspective are those of Amihud and Mendelson (1986), Constantinides (1986), Heaton and Lucas (1996), Vayanos (1998), Huang (2003), and Lo et al. (2004), among others.

The model of Acharya and Pedersen (2005) is part of the first as well as the second category. The authors develop a theoretical equilibrium model (a liquidity-adjusted CAPM) that considers both the levels of asset liquidity and their covariance matrix. The resulting pricing equation relates the expected excess return on any asset to its expected level of liquidity, as well as to four betas. These betas measure the covariances between (1) the asset's return and the market return, (2) the asset's illiquidity and market illiquidity, (3) the asset's return and market illiquidity, and (4) the asset's illiquidity with the market return. The authors test their model using NYSE and AMEX stocks over the sample period 1963 to 1999 and employ the illiquidity measure proposed by Amihud (2002). They conclude that their liquidity-adjusted CAPM does a better job than the CAPM in explaining the cross section of liquidity-, liquidity variation-, and size-sorted portfolios. However, it seems that their model is unable to capture the value effect.

Two further models that fall into the second category are those developed by Pástor and Stambaugh (2003) and Liu (2006). Liu (2006) suggests a two-factor model to describe expected excess returns. The first factor is the equity premium. The second factor is the difference between the returns on diversified portfolios of low- and high-liquidity stocks. Liu (2006) tests this model using NYSE, AMEX, and NASDAQ stocks over the period 1960 to 2003. The author concludes that—in contrast to the CAPM and FF3— the model does a good job explaining the returns on portfolios of stocks sorted by liquidity risk (i.e., by their betas on the liquidity factor). Moreover, Liu concludes that the two-factor model is able to explain CAPM anomalies associated with size; the book-to-market, cash flow-to-price, earnings-to-price, and dividend-to-price ratios; and long-term reversal.

Pástor and Stambaugh (2003) use an equally weighted average of the liquidity measures of individual stocks to construct a measure of market liquidity in a given month and their liquidity risk factor is the innovation to market liquidity. The monthly individual liquidity measures employ daily data within that month. In particular, the liquidity measure for stock i in month t, $\hat{l}_{i,t}$, is computed by estimating $l_{i,t}$ in the following regression:

$$R^{em}_{i,d+1,t} = \theta_{i,t} + \phi_{i,t}R_{i,d,t} + l_{i,t}sign\left(R^{em}_{i,d,t}\right) \cdot \mathrm{V}_{i,d,t} + \varepsilon_{i,d+1,t}, \quad d = 1, \dots, D, \tag{2.13}$$

where $R_{i,d,t}$ is the return on stock i on day d in month t, $R^{em}_{i,d,t}$ is the difference between $R_{i,d,t}$ and the market return on day d in month t, $\mathrm{V}_{i,d,t}$ is the US dollar volume for stock i on day d in month t, and $\varepsilon_{i,d+1,t}$ is a disturbance term. The regression is only conducted if there are more than 15 observations, that is, $D > 15$.[8] A stock's liquidity in month t measured that way can be interpreted as the average impact that a given volume on day d has on the return for day $d + 1$, where the volume has the same sign as the return on day d. Roughly speaking, higher illiquidity is related to stronger volume-related return reversals. On this basis, Pástor and Stambaugh compute their marketwide measure of liquidity as

$$\hat{l}_t = \frac{1}{N_t}\sum_{i=1}^{N_t} \hat{l}_{i,t}, \tag{2.14}$$

where N_t denotes the number of stocks for which Pástor and Stambaugh obtain an estimate of $l_{i,t}$ in month t. Afterward, the authors compute scaled differences in these monthly liquidity measures:

$$\Delta\hat{l}_t = \left(\frac{m_t}{m_1}\right)\frac{1}{N_t}\sum_{i=1}^{N_t}\left(\hat{l}_{i,t} - \hat{l}_{i,t-1}\right), \tag{2.15}$$

where m_t represents the total US dollar value at the end of month $t - 1$ of the stocks that are used to compute $\hat{l}_t$ in month t and month $t = 1$ (for m_1) is August 1962. Then, Pástor and Stambaugh conduct regressions

$$\Delta\hat{l}_t = a + b\Delta\hat{l}_{t-1} + c\left(\frac{m_{t-1}}{m_1}\right)\hat{l}_{t-1} + u_t \tag{2.16}$$

to estimate fitted residuals in marketwide liquidity, $\hat{u}_t$. Finally, the innovation in aggregate liquidity (i.e., the liquidity factor) is obtained as

$$L_t = \frac{1}{100}\hat{u}_t. \tag{2.17}$$

On this basis, Pástor and Stambaugh (2003) propose an asset pricing model (PS) that augments FF3 with L_t. The corresponding time-series regression is

$$R_{i,t} - R_{f,t} = \alpha_i + b_i\left(R_{m,t} - R_{f,t}\right) + s_i SMB_t + h_i HML_t + \beta_{iL}L_t + e_{i,t}, \tag{2.18}$$

where β_{iL} is the liquidity beta of stock i. Pástor and Stambaugh estimate their model over the sample period 1962 to 1999, using NYSE, AMEX, and NASDAQ stocks. They construct 10 portfolios in which stocks are sorted by their predicted or historical liquidity betas (LIQ10).

[8] Pástor and Stambaugh (2003) provide a few more technical details.

Then, the authors use a time-series regression approach to test whether the average excess returns on these portfolios can be explained by the CAPM, FF3, and C. They conclude that none of these models can (completely) explain these portfolios. While the alphas of the portfolios with low-liquidity betas tend to be significantly negative, the alphas of portfolios with high-liquidity betas are typically significantly positive. Consequently, Pástor and Stambaugh conclude that stocks with higher-liquidity betas, that is, stocks that are more sensitive to variation in aggregate liquidity, have higher expected returns and that this fact cannot be explained by the stocks' covariations with the equity premium, Fama and French's (1993) size and value factors, or Carhart's (1997) momentum factor.

Figure 5 plots the average excess returns on the LIQ10 portfolios against their market betas using updated data from January 1968 through December 2012. The portfolios plotted are sorted on historical liquidity betas (Pástor and Stambaugh, 2003, Sec. III.B). Note that there is much dispersion in average excess returns that is unrelated or even negatively related to market beta, which indicates that this variation cannot be explained by the CAPM. In particular, the spread in average excess returns of the portfolio that consists of the stocks with the highest-liquidity betas (denoted by 10 in Figure 5) and the portfolio of stocks that show the lowest-liquidity betas (denoted by 1 in Figure 5) is, at around 0.45% per month, economically significant.

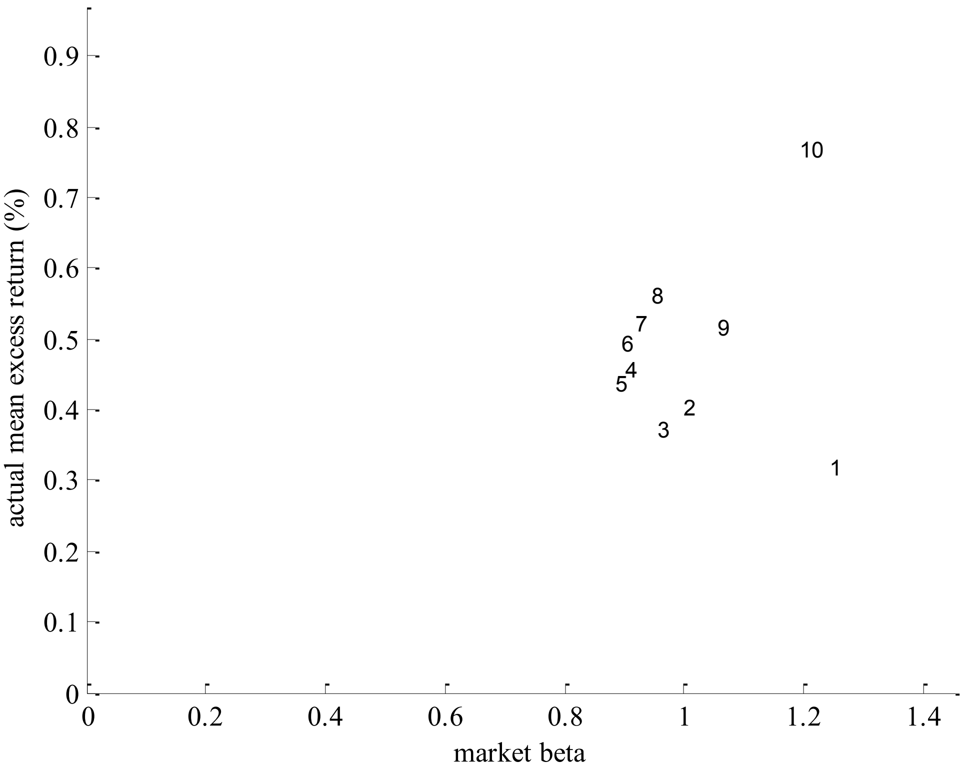

Figure 5. Mean excess returns versus market betas (LIQ10).
This figure plots the market betas of 10 value-weighted portfolios of NYSE, AMEX, and NASDAQ stocks sorted by historical liquidity betas against their percentage monthly mean returns in excess of the one-month US Treasury bill rate, measured over the sample period January 1968 to December 2012. In this figure, 1 denotes the portfolio that consists of the stocks with the lowest-liquidity betas and 10 denotes the portfolio of the stocks with the highest-liquidity betas. The notation of the remaining portfolios follows this scheme. I thank Lubos Pástor for providing me with the return series for these 10 portfolios. The US Treasury bill rate data are obtained from Kenneth R. French's website. A detailed description of the construction of these 10 portfolios is given in Sec. III.B of Pástor and Stambaugh (2003).

2.2.1.6. Empirical applications of the ICAPM

Many recent attempts seek to link CAPM anomalies discovered in the cross section of expected returns to the observations of time-varying expected returns (Sec. 2.1.1). The theoretical framework of the ICAPM seems to be ideally suited for this project. The ICAPM predicts that variables that exhibit forecasting power for the distribution of future returns should be priced in the cross section of returns. These additional factors may be able to explain anomalous effects such as size, value, and momentum. The ICAPM remained a purely theoretical model for the first 20 years after its publication mainly because Merton (1973) does not directly identify the state variables that predict consumption investment opportunities, although the author suggests employing the Treasury bill rate as such a variable. But with the accumulated empirical evidence that future returns—and hence investment opportunities—are predictable by various factors, researchers seemed to finally recognize potential Mertonian state variables (Cochrane, 2005, Sec. 9.3). Consequently, more than 20 years after its publication, researchers started to propose empirical applications of the

ICAPM. These models try out various state variables that are potential predictors of investment opportunities as factors in addition to the market. Additionally, researchers use the ICAPM to interpret the success of empirically motivated multifactor models such as FF3. Hence, the ICAPM has become the theory behind many multifactor models. Fama (1991, p. 1594) emphasizes that the ICAPM (as well as the APT) provides a great opportunity for empirical researchers by justifying asset pricing models with multiple factors: “The multifactor models are an empiricist’s dream. They are off-the-shelf theories that can accommodate tests for cross-sectional relations between expected returns and the loadings of security returns on any set of factors that are correlated with returns.”

The first attempts to use predictive state variables within a multifactor model were those of Chen et al. (1986) and Ferson and Harvey (1991). They employ factors such as the growth rate of industrial production, the default yield spread, and the term structure spread. Both models are justified as possible empirical applications of the ICAPM. Additionally, Shanken (1990) proposes an ICAPM that employs the Treasury bill rate and a measure of Treasury bill rate volatility as state variables.

More recently, Brennan et al. (2004) include the real interest rate and the maximum Sharpe ratio as state variables in their empirical application of the ICAPM. The authors conclude that their model outperforms both the CAPM and FF3 in explaining the 25 size/book-to-market portfolios proposed by Fama and French (1993). Ang et al. (2006) provide evidence that innovations in aggregate volatility are negatively priced in the cross section of average stock returns. Since aggregate volatility seems to also predict future market returns (see Sec. 2.1.1), this finding is consistent with aggregate volatility being a state variable in the sense of Merton’s theory. Gerard and Wu (2006) suggest a two-factor ICAPM that includes the long-term interest rate as a state variable.

Lo and Wang (2006) propose a two-factor ICAPM that comprises the returns on the market portfolio and on a portfolio that hedges shifts in consumption investment opportunities. This hedging portfolio is identified using the return and trading volume data of individual stocks. The authors show that their hedging portfolio predicts future market returns (i.e., investment opportunities) and they find that their two-factor model does reasonably well in explaining the cross-sectional variation of expected returns. Guo and Savickas (2008) show that average idiosyncratic volatility forecasts future stock market returns and that this factor helps to explain the cross section of Fama and French’s (1993) 25 size/book-to-market portfolios just as well as FF3’s value factor. These results support the ICAPM theory and indicate that average idiosyncratic volatility is a Mertonian state variable. Ozoguz (2009) proposes an ICAPM that employs investors’ conditional beliefs and their uncertainty about the state of the economy as state variables. These factors are measured using a regime-switching model with market returns and aggregate output. Further empirical applications of the ICAPM are proposed by Hahn and Lee (2006) and Petkova (2006), which are presented in more detail later in this section.

Beyond these papers, there is one body of literature that investigates variants or extensions of the ICAPM suggested by Campbell (1993). These papers include those of Campbell (1996), Campbell and Vuolteenaho (2004), Guo (2006b), Chen and Zhao (2009), and Campbell et al. (2013).

Campbell and Vuolteenaho (2004) argue that the return on the market portfolio comprises two components, cash flow news and discount rate news, and propose a two-factor ICAPM. Specifically, they state that the market may fall because of bad news about future cash flows or because of an increase in the discount rate applies to these cash flows. However, bad cash flow news decreases the wealth of investors without changing investment opportunities, while bad discount rate news (i.e., a rise in the discount rate) decreases wealth, on the one hand, but improves investment opportunities, on the other hand. Consequently, a risk-averse long-term investor may require an asset that comoves with the market's cash flow news to provide a higher premium on its expected return than an asset that covaries with the market's discount rate news, since the decrease in wealth resulting from the higher discount rate is to some extent compensated by the improved investment opportunities, that is, the higher expectations with respect to future returns. The authors state that taking into account the difference between these two components can explain the size, value, and momentum anomalies. For this reason, Campbell and Vuolteenaho break the market beta into two different betas: cash flow beta and discount rate beta, or "bad beta" and "good beta," respectively. They state that, according to the ICAPM, the discount rate beta requires a risk price that corresponds to the variance of the market return, while the cash flow beta demands a risk price that is γ times greater, where γ denotes the coefficient of the relative risk aversion of the representative investor (RRA).

To be specific, Campbell and Vuolteenaho (2004) follow the return decomposition framework of Campbell and Shiller (1988a) and Campbell (1991) to estimate cash flow news and discount rate news. First, they assume that the vector of state variables, z_{t+1}, follows a first-order vector autoregressive (VAR) process:

$$z_{t+1} = a + \Gamma z_t + u_{t+1}, \tag{2.19}$$

where a and Γ denote a vector and matrix of constant parameters, respectively, and u_{t+1} is an i.i.d. vector of shocks. Then, the cash flow news and discount rate news are formulated as linear functions of u_{t+1}:

$$\begin{gathered} N_{CF,t+1} = (e1' + e1'\lambda)u_{t+1}, \\ N_{DR,t+1} = e1'\lambda u_{t+1}, \end{gathered} \tag{2.20}$$

where $N_{CF,t+1}$ and $N_{DR,t+1}$ represent the cash flow and discount rate news, respectively, at time $t+1$. Moreover, $e1$ is a standard basis vector with one as its first element and zeros as its remaining elements. Furthermore, λ is defined as $\lambda \equiv \rho\Gamma(I - \rho\Gamma)^{-1}$, where ρ is a discount coefficient and I denotes the identity matrix. Campbell and Vuolteenaho set the log excess market return in t as the first element of z_t. The remaining elements are the term structure spread, the aggregate price–earnings ratio, and the small-stock value spread.

In the next step, Campbell and Vuolteenaho (2004) define the cash flow beta of asset i, β_{iCF}, as the covariance of the asset's return with cash flow news, divided by the variance of unexpected market returns,

$$\beta_{iCF} \equiv \frac{Cov(r_{i,t}, N_{CF,t})}{Var[r_{m,t}^{e} - E_{t-1}(r_{m,t}^{e})]}, \tag{2.21}$$

and they define the asset's discount rate beta as the ratio of the covariance of the asset's return with *good* discount rate news (i.e., lower than expected discount rates) to the variance of unexpected market returns:

$$\beta_{iDR} \equiv \frac{Cov(r_{i,t}, -N_{DR,t})}{Var[r_{m,t}^{e} - E_{t-1}(r_{m,t}^{e})]}, \tag{2.22}$$

where $r_{i,t}$ is the log return on asset i at time t and $r_{m,t}^{e}$ is the excess log return on the market. Campbell and Vuolteenaho note that β_{iCF} and β_{iDR} add up to the total market beta, $\beta_{im} = \beta_{iCF} + \beta_{iDR}$. On this basis, Campbell and Vuolteenaho formulate their two-beta ICAPM as follows (using simple returns):

$$E(R_i - R_f) = \gamma \sigma_m^2 \beta_{iCF} + \sigma_m^2 \beta_{iDR}, \tag{2.23}$$

where σ_m^2 is the variance of the return on the market portfolio. The authors test Eq. (2.23) as well as the CAPM and an unrestricted version of their two-beta ICAPM. I denote the latter as CV hereafter. In particular, the authors carry out the cross-sectional regression

$$\overline{R_i - R_f} = g_0 + g_1 \hat{\beta}_{iCF} + g_2 \hat{\beta}_{iDR} + e_i, \tag{2.24}$$

where $\overline{R_i - R_f}$ is the sample average simple excess return on asset i, $\hat{\beta}_{iCF}$ and $\hat{\beta}_{iDR}$ are estimates for β_{iCF} and β_{iDR}, respectively, and e_i is a disturbance term (the pricing error). To test the two-factor ICAPM in Eq. (2.23), g_2 is restricted to the estimated variance of the market return, the CAPM is estimated by restricting g_1 to equal g_2, and the CV model (the unrestricted version) allows unrestricted values for g_1 and g_2, that is, free risk prices for cash flow and discount rate betas, respectively. Campbell and Vuolteenaho estimate two different specifications of these three models. The first one assumes the absence of a risk-free rate and includes an unrestricted zero-beta rate, R_z, following Black (1972). The second one assumes that the zero-beta rate equals R_f. To estimate the second specification, the intercept g_0, which represents the difference between R_z and R_f, is restricted to zero (i.e., excluded from the regression).

Campbell and Vuolteenaho (2004) employ 45 test assets, namely, SBM25 and 20 portfolios sorted by market beta, and consider the two sample periods 1929 to 1963 and 1963 to 2001. The authors find that the CAPM, the two-beta ICAPM, and CV all reasonably explain the cross section of average returns in the early sample. They conclude that the rather good performance of the CAPM is due to the fact that the ratio of the discount rate beta to the cash flow beta is relatively constant across assets during this period. However, with regard to the more recent sample, "the CAPM fails disastrously to explain the returns on the test assets" (Campbell and Vuolteenaho, 2004, p. 1265). The poor performance of the CAPM is greatly improved by the other two models. The reason seems to be that the high betas of growth stocks (which have low average returns) are predominantly discount rate betas (which demand low risk prices), while value stocks (which have high average returns) have higher cash flow betas (which require higher risk premiums). Moreover, the flat relation between the market beta and average returns (which is observed in the more recent CAPM tests; see Sec. 2.2.1.3) can be explained by the fact that stocks with high market betas have higher discount rate betas but almost the same cash flow betas as stocks with lower market betas.

Petkova (2006) assumes that the unconditional expected return on any asset i is described by the following general model:

$$E(R_i - R_f) = \lambda\beta_{im} + \sum_{k=1}^{K} \lambda_k \beta_{ik}, \tag{2.25}$$

where λ denotes the equity premium and λ_k is the beta risk price associated with innovations in state variable k. The betas, β_{im} and β_{ik}, are the slope coefficients from the return-generating process

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_{im}(R_{m,t} - R_{f,t}) + \sum_{k=1}^{K} \beta_{ik} u_{k,t} + \varepsilon_{i,t}, \tag{2.26}$$

where $u_{k,t}$ is the innovation to state variable k at time t. The innovation is the unexpected change of the state variable and, according to the asset pricing model, only this component should require a risk premium. Moreover, these innovations and the excess return on the market are realized contemporaneously. Petkova specifies a VAR process to derive the state variable innovations, following Campbell (1996). In particular, Petkova defines a state vector z_t, whose first element is the market excess return in t. The remaining elements of z_t are the values of the K state variables in t. Petkova assumes that the demeaned vector z_t follows a first-order VAR process:

$$z_t = Az_{t-1} + u_t, \tag{2.27}$$

where A is a matrix of constant parameters (the slope coefficients) and u_t is a vector of residual values at time t. The element in u_t that corresponds to state variable k is the innovation to k in Eq. (2.26).

Petkova (2006) employs the short-term risk-free rate, the term structure spread, the dividend yield, the default yield spread, as well as HML and SMB as state variables. The author emphasizes that the first four variables are known to have forecasting power for investment opportunities. On this basis, Petkova estimates three models: a seven-factor model that includes the innovations to all six state variables, a five-factor model that incorporates only the innovations to the first four variables (where the VAR system used to obtain the innovations still includes HML and SMB), and FF3. I denote the five-factor specification as P hereafter. The testing assets are SBM25 and the sample period is July 1963 to December 2001. Petkova shows, among other things, that the innovations to HML and SMB do not provide any additional explanatory power for the cross section of average returns within the seven-factor model. Hence, HML and SMB seem to lose their information content when confronted with predictors of investment opportunities. Moreover, the author concludes that P exhibits a higher explanatory power than FF3.

Hahn and Lee (2006) propose another application of the ICAPM, namely, HL. Their model is very similar to P. The main difference is that Hahn and Lee only employ the term structure spread and default yield spread as state variables. Moreover, they do not estimate a VAR system to derive the state variable innovations. Instead, they use first differences to proxy for them. In particular, the innovation to the term structure spread at time t is defined as

$$\Delta TERM_t \equiv TERM_t - TERM_{t-1}, \tag{2.28}$$

where $TERM_t$ denotes the term structure spread at time t. The innovation to the default yield spread, ΔDEF, is computed similarly, except that the authors use the negative of the change, to obtain a positive correlation between ΔDEF and SMB.[9] The authors state that a simple autoregressive specification for the state variables leads to almost identical results. Hahn and Lee use SBM25 as testing assets and start by running the following time-series regression for each portfolio i over the sample period July 1963 to June 2001:

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_{im}\left(R_{m,t} - R_{f,t}\right) + \beta_{iDEF}\Delta DEF_t + \beta_{iTERM}\Delta TERM_t + e_{i,t}. \tag{2.29}$$

Afterward, they use the estimated betas, $\hat{\beta}_{im}$, $\hat{\beta}_{iDEF}$, and $\hat{\beta}_{iTERM}$, to conduct month-by-month Fama–MacBeth (1973) cross-sectional regressions

$$R_{i,t} - R_{f,t} = \lambda_{0,t} + \lambda_t\hat{\beta}_{im} + \lambda_{DEF,t}\hat{\beta}_{iDEF} + \lambda_{TERM,t}\hat{\beta}_{iTERM} + \varepsilon_{i,t}, \tag{2.30}$$

where $\lambda_{0,t}$ is the constant and λ_t, $\lambda_{DEF,t}$, and $\lambda_{TERM,t}$ are the slope coefficients of the cross-sectional regression in month t. The equity premium as well as the beta risk prices associated with DEF and $TERM$ are then obtained as the means of the monthly estimates of λ_t, $\lambda_{DEF,t}$, and $\lambda_{TERM,t}$. Hahn and Lee estimate FF3 analogously and compare their estimation results. They conclude that ΔDEF and $\Delta TERM$ capture most of the explanatory power provided by SMB and HML and that SMB and HML become superfluous in the presence of ΔDEF and $\Delta TERM$. In particular, the betas on ΔDEF and SMB vary similarly across portfolios along the size dimension of SBM25 and the loadings on $\Delta TERM$ and HML show the same behavior along the book-to-market dimension. Moreover, the risk premiums associated with ΔDEF and $\Delta TERM$ seem to be similar to those associated with SMB and HML.

2.2.1.7. Multifactor models and their consistency with the ICAPM

The ICAPM places restrictions that must be satisfied by a multifactor model to be justifiable by Merton's theory, as it is emphasized by MSC. These authors empirically study FF3, FF5 (Sec. 2.2.1.3), C (Sec. 2.2.1.4), PS (Sec. 2.2.1.5), and CV, P, and HL (Sec. 2.2.1.6). Beyond that, they examine the model proposed by Koijen et al. (2010), hereafter KLVN, which incorporates three factors: the equity premium, as well as innovations to the term structure spread and to the Cochrane–Piazzesi factor, presented in Sec. 2.1.2. According to MSC (p. 586), these multifactor models "represent some of the most relevant examples presented in the empirical asset pricing literature." They conclude that only FF3 and C can be justified by the ICAPM.

[9] When I test HL later in this thesis, I do not use the negative of the change, following MSC.

MSC start their analysis by presenting a simple version of the ICAPM that is based on a representative investor's consumption/portfolio choice problem in continuous time.[10] Pennacchi (2008, Ch. 13.1) describes a similar specification in more detail. In this model, both the mean and volatility of asset returns are functions of a single state variable, z, which evolves as a diffusion process through time, so that investment opportunities are time varying. The authors state that the model's equilibrium relation between expected return and risk can be approximated in discrete time as

$$E_t(R_{i,t+1}) - R_{f,t+1} = \gamma Cov_t(R_{i,t+1}, R_{m,t+1}) + \gamma_z Cov_t(R_{i,t+1}, \Delta z_{t+1}), \tag{2.31}$$

where $R_{i,t+1}$ denotes the return on asset i between time t and $t+1$, $R_{f,t+1}$ is the risk-free rate known at t, and $R_{m,t+1}$ is the return on the market portfolio (see also Cochrane, 2005, Sec. 9.2, and Maio, 2013). In addition, $Cov_t(\cdot)$ denotes the conditional covariance and γ denotes the RRA. Furthermore, Δz_{t+1} is the innovation or change between times t and $t+1$ to the state variable z and γ_z represents the (covariance) risk price associated with the state variable. The term $\gamma_z Cov_t(R_{i,t+1}, \Delta z_{t+1})$ distinguishes the pricing equation from that of the CAPM. Through the law of iterated expectations, MSC obtain an ICAPM pricing equation in unconditional form:

$$E(R_{i,t+1} - R_{f,t+1}) = \gamma Cov(R_{i,t+1}, R_{m,t+1}) + \gamma_z Cov(R_{i,t+1}, \Delta z_{t+1}). \tag{2.32}$$

On this basis, MSC derive several restrictions that the ICAPM places on the time-series and cross-sectional behavior of the state variables. Specifically, if a multifactor model is justified as an empirical application of the ICAPM, it must satisfy the following criteria.

ICAPM CRITERION 1: *The market (covariance) price of risk must be economically plausible as an estimate of the RRA and thus have a value between one and 10.*

The first ICAPM criterion is associated with the theoretical postulation that if an asset's covariance with the market return is positive, it earns a positive risk premium over the risk-free rate, since a risk-averse investor cannot use such an asset to hedge the risk of changes in current aggregate wealth. In particular, in an empirical application of the ICAPM, the estimate of γ must be economically plausible as an estimate of the RRA. According to, for example, Mehra and Prescott (1985), an economically plausible estimate would be between one and 10.

ICAPM CRITERION 2: *The candidate state variables must forecast expected market returns or market volatility.*

The state variable z is related to changes in the investor's set of investment opportunities. This implies that it must forecast the distribution of future aggregate returns, that is, their first or second moments, which becomes the second ICAPM criterion.

The third ICAPM criterion derived by MSC is associated with an investor's reinvestment risk, captured by $\gamma_z Cov(R_{i,t+1}, \Delta z_{t+1})$. I divide this third criterion into criteria 3a and 3b. Criterion 3a considers changes in the set of investment opportunities that are driven by changes in

[10] The following presentation is based on Section 2 of Lutzenberger (2014b).

expected market returns, whereas criterion 3b regards changing investment opportunities that are represented by changes in market volatility.

ICAPM CRITERION 3A: *If a state variable positively (negatively) forecasts expected market returns in the time series, its innovation should earn a positive (negative) risk price in the cross section.*

If the state variable z_{t+1} (and thus its innovation Δz_{t+1}) covaries positively with the future (expected) market return,

$$\begin{aligned} &Cov_t(R_{m,t+2}, z_{t+1}) \\ &= Cov_t[E_{t+1}(R_{m,t+2}), z_{t+1}] = Cov_t[E_{t+1}(R_{m,t+2}), \Delta z_{t+1}] > 0, \end{aligned} \tag{2.33}$$

and if, at the same time, the return on asset i covaries, as an assumption, positively with the (innovation in the) state variable,

$$Cov_t(R_{i,t+1}, z_{t+1}) = Cov_t(R_{i,t+1}, \Delta z_{t+1}) > 0, \tag{2.34}$$

then the covariance between the return on asset i and the future (expected) market return is also positive:

$$Cov_t(R_{i,t+1}, R_{m,t+2}) = Cov_t[R_{i,t+1,}E_{t+1}(R_{m,t+2})] > 0. \tag{2.35}$$

Such assets cannot be used by investors to hedge their reinvestment risk, since these assets provide low returns when the expected future market return (i.e., investors' future investment opportunities) is also low. Therefore, a rational investor requires such an asset to offer a higher risk premium than an asset that is not correlated with the future market return; that is, the risk premium associated with the state variable must be positive, $\gamma_z Cov(R_{i,t+1}, \Delta z_{t+1}) > 0$. As an implication, given the assumption $Cov_t(R_{i,t+1}, \Delta z_{t+1}) > 0$, the risk price for intertemporal risk is positive, $\gamma_z > 0$. If, on the other hand, the state variable is negatively correlated with future market returns, γ_z must be negative.

ICAPM CRITERION 3B: *If a state variable positively (negatively) forecasts the volatility of market returns in the time series, its innovation should earn a negative (positive) risk price in the cross section.*

If the state variable z_{t+1} (and thus its innovation Δz_{t+1}) covaries positively with the future volatility of the market return,

$$\begin{aligned} &Cov_t(R_{m,t+2}^2, z_{t+1}) \\ &= Cov_t[E_{t+1}(R_{m,t+2}^2), z_{t+1}] = Cov_t[E_{t+1}(R_{m,t+2}^2), \Delta z_{t+1}] > 0, \end{aligned} \tag{2.36}$$

and if the return on asset i still covaries, as an assumption, positively with the (innovation in the) state variable, then the covariance between the return on asset i and the future volatility of the market return is also positive:

$$Cov_t(R_{i,t+1}, R_{m,t+2}^2) = Cov_t[R_{i,t+1}, E_{t+1}(R_{m,t+2}^2)] > 0. \tag{2.37}$$

Such an asset provides a risk-averse investor (who dislikes volatility) a hedge for reinvestment risk, since it creates high returns in times when future market volatility is also high. As a result, an investor should require such an asset to offer a lower risk premium than an asset that does not covary with future market volatility, that is, $\gamma_z Cov(R_{i,t+1}, \Delta z_{t+1}) < 0$. As an implication, given the assumption $Cov_t(R_{i,t+1}, \Delta z_{t+1}) > 0$, the risk price for intertemporal risk must be negative, $\gamma_z < 0$. If the state variable is, instead, assumed to be negatively correlated with future market volatility, then γ_z must be positive.

MSC categorize the eight multifactor models to test into two groups. The first group includes models that are explicitly justified as ICAPM applications and consists of HL, P, CV, and KLVN.[11] As risk factors in addition to the market equity premium, these models use innovations in state variables that are known from the predictability of returns literature, in which they are used to forecast market returns in the time series. The second category comprises multifactor models that are less justified by the ICAPM but, rather, empirically motivated or based on liquidity risk. However, some authors consider them possible applications of the ICAPM. This category includes FF3, C, PS, and FF5.

MSC formulate each model in an expected return–covariance representation. For instance, they formulate CV as

$$\begin{aligned} E(R_{i,t+1} - R_{f,t+1}) \\ &= \gamma Cov(R_{i,t+1} - R_{f,t+1}, RM_{t+1}) \\ &+ \gamma_{TERM} Cov(R_{i,t+1} - R_{f,t+1}, \Delta TERM_{t+1}) \\ &+ \gamma_{PE} Cov(R_{i,t+1} - R_{f,t+1}, \Delta PE_{t+1}) \\ &+ \gamma_{VS} Cov(R_{i,t+1} - R_{f,t+1}, \Delta VS_{t+1}), \end{aligned} \tag{2.38}$$

where RM_{t+1} denotes the excess market return in $t+1$ and $\Delta TERM_{t+1}$, ΔPE_{t+1}, and ΔVS_{t+1} represent the innovations in $t+1$ to the term structure spread, the aggregate price-earnings ratio, and the small-stock value spread, respectively. Moreover, γ_{TERM}, γ_{PE}, and γ_{VS} are the (covariance) risk prices associated with these three state variables. MSC state that this specification of CV is equivalent to that presented by Campbell and Vuolteenaho (2004), which includes only two factors (cash flow news and discount rate news), since they are linear functions of the innovations in the equity premium and the three state variables.

The innovations in the state variables are, as in the work of Hahn and Lee (2006), proxied by first differences. The remaining models are formulated similarly and all formulas are documented by MSC. Therefore and to save space, I skip the presentation of each formula and instead summarize the remaining models in Table 2.

[11] Koijen et al. (2010) do not actually justify KLVN as an empirical application of the ICAPM. Nevertheless, MSC include the model in this group, since it incorporates two variables from the predictability of returns literature and, hence, is constructed very similarly to HL, P, and CV.

Table 2. Multifactor models investigated by MSC.
This table provides an overview of the multifactor models included by MSC. The first column contains the reference and abbreviation for the model and the second column presents the factors included in the model.

Model	Factors
Hahn and Lee, 2006 (HL)	Excess market return, RM Innovation in term structure spread, $\Delta TERM$ Innovation in default yield spread, ΔDEF
Petkova, 2006 (P)	Excess market return, RM Innovation in term structure spread, $\Delta TERM$ Innovation in default yield spread, ΔDEF Innovation in market dividend–price ratio, ΔDY Innovation in short-term risk-free rate, ΔRF
Campbell and Vuolteenaho, 2004 (CV)	Excess market return, RM Innovation in term structure spread, $\Delta TERM$ Innovation in aggregate price-earnings ratio, ΔPE Innovation in small-stock value spread, ΔVS
Koijen et al., 2010 (KLVN)	Excess market return, RM Innovation in term structure spread, $\Delta TERM$ Innovation in the Cochrane–Piazzesi factor, ΔCP
Fama and French, 1993 (FF3)	Excess market return, RM Size factor (small minus big), SMB Value factor (high minus low), HML
Fama and French, 1993 (FF5)	Excess market return, RM Size factor (small minus big), SMB Value factor (high minus low), HML Innovation in term structure spread, $\Delta TERM$ Innovation in default yield spread, ΔDEF
Carhart, 1997 (C)	Excess market return, RM Size factor (small minus big), SMB Value factor (high minus low), HML Momentum factor, UMD
Pástor and Stambaugh, 2003 (PS)	Excess market return, RM Size factor (small minus big), SMB Value factor (high minus low), HML Liquidity-related risk factor, L

Note that MSC construct FF5 using the innovation in the term structure spread and the innovation in the default yield spread (which are both based on yields), while Fama and French (1993) employ, as shown in Eq. (2.11), the long-term return spread and the default return spread (which are constructed with returns rather than yields). I follow MSC and use the two yield-based factors, when I later estimate the model in the European stock market. MSC use SBM25 as well as SM25 as testing assets. Their sample period is July 1963 to December 2008. Table 3 summarizes their results.

Table 3. Consistency of multifactor models with the ICAPM.

This table shows whether MSC's ICAPM criteria are satisfied in the US stock market (MSC, Table 1). A check mark (✓) means that the respective criterion is satisfied. ICAPM criterion 2 is satisfied by all models and is independent of the testing assets. Therefore, it is not displayed in this table. Panel A displays the results for 25 portfolios sorted by size and book-to-market ratio as test assets and Panel B presents the results for 25 portfolios sorted by size and momentum. The sample period is July 1963 to December 2008.

	ICAPM criterion 1	ICAPM criterion 3a	ICAPM criterion 3b
Panel A: SBM25			
HL	x	x	x
P	x	x	x
CV	x	x	x
KLVN	x	✓	x
FF3	✓	✓	✓
C	✓	✓	x
PS	x	✓	x
FF5	x	x	✓
Panel B: SM25			
HL	✓	x	x
P	x	x	x
CV	x	x	x
KLVN	x	x	x
FF3	x	x	x
C	✓	✓	x
PS	x	x	x
FF5	x	x	x

Note that when the testing assets are SBM25 and shifts in investment opportunities are associated with expected market returns, both FF3 and C can be justified by the ICAPM, since they meet the ICAPM criteria 1, 2, and 3a. If MSC regard changes in investment opportunities driven by market volatility, only FF3 is consistent with the ICAPM, since the model meets the ICAPM criteria 1, 2, and 3b. Moreover, note that if MSC use SM25 as testing assets and consider shifts in investment opportunities that are associated with expected market returns, only C is justifiable as an empirical application of the ICAPM. If one assumes changes in investment opportunities driven by market volatility, however, none of the models investigated meets the ICAPM criteria.

2.2.2. International stock returns and other asset classes

It took some time before researchers started testing the CAPM and other factor models with data from outside the US. One of the first studies using non-US data is that of Chan et al. (1991), who find that the book-to-market ratio has strong explanatory power for the average returns of Japanese stocks, controlling for market beta. Haugen and Baker (1996) examine the returns of US, French, German, Japanese, and UK stocks and conclude that the characteristics that determine expected returns are quite stable from country to country. They find that classic risk measures such as market beta and volatility do not have any robust influence on expected returns, while the most robust determinants are past returns, trading volume, and financial ratios such as return on equity and the price–earnings ratio. Fama and French (1998) find that portfolios of stocks with high book-to-market ratios produce higher average returns than portfolios of stocks with low book-to-market ratios in 12 of 13 countries and a global portfolio that buys high–book-to-market stocks and sells low–book-to-market stocks from all 13 countries produces an average return of 7.68% per year. They obtain similar value premiums for portfolios sorted on earnings-to-price, cash flow-to-price, and dividend-to-price ratios. This value effect cannot be explained by a global CAPM whose market portfolio comprises all 13 countries at once. A variety of other papers investigate whether expected returns are better described globally or locally. Reviews are provided by Karolyi and Stulz (2003) and Lewis (2011). Griffin (2002) investigates country-specific versus global versions of FF3 and finds that domestic models do a better job explaining the time-series variation in international stock returns and exhibit lower pricing errors. The author concludes that applications of FF3 should be conducted on a country-specific basis. Moreover, Rouwenhorst (1998) provides evidence on the momentum effect in Europe, Liew and Vassalou (2000) find international evidence on both value and momentum, and Griffin et al. (2003) and Chui et al. (2010) provide further insights on momentum in international stock markets. Furthermore, Hou et al. (2011) find that a global three-factor model that comprises a global cash flow-to-price factor, a global momentum factor, as well as a global market factor captures much time variation in global stock returns. Beyond that, Lee (2011) examines a global version of the liquidity-adjusted CAPM of Acharya and Pedersen (2005), while Li et al. (2014) test Amihud's (2002) illiquidity measure and the liquidity-adjusted CAPM of Acharya and Pedersen (2005) over Japanese stock returns.

Ziegler et al. (2007) examine FF3 and FF5 within the German stock market. Using time-series regression tests, they conclude that FF3 has greater explanatory power than the CAPM for the average excess returns on 16 portfolios of German stocks sorted by size and book-to-market ratio. However, FF5 does not provide additional explanatory power to FF3. Bauer et al. (2010) examine FF3 (as well as a dynamic version of it) within the European stock market. Among other things, they find a significant size effect, which is in contrast to the US evidence, that the size effect has vanished since its discovery. Similar to the US evidence, the GRS test on Bauer et al.'s European version of the SBM25 rejects the null hypothesis that the alphas produced by FF3 are jointly equal to zero. However, the small–growth portfolio shows a significantly positive alpha. Hence, this portfolio seems to have an expected excess return that is too *high* to be explained by FF3, a result that disagrees with the US evidence of Fama and French (1993) that this portfolio exhibits a pricing error that is significantly negative.

Artmann et al. (2012a) find that average returns on German stocks are, according to portfolios sorted by a single characteristic, related to the book-to-market ratio, the earnings-to-price ratio, leverage, return on assets, and momentum. However, their Fama–MacBeth (1973) regressions suggest that only the book-to-market ratio, the earnings-to-price ratio, and momentum have a significant impact on average returns. Additionally, the authors show that FF3 performs rather poorly in Germany and is outperformed by C, using a variety of double-sorted portfolios as testing assets. Employing a test proposed by Patton and Timmermann (2010), Artmann et al. (2012b) conclude that average returns across 10 portfolios of German stocks sorted on momentum increase monotonically, while such a monotonic relation does not exist for portfolios sorted on beta, size, or the book-to-market ratio, respectively. Moreover, the authors conclude that the CAPM, FF3, and C do a good job explaining the average returns on portfolios of German stocks sorted by beta or industry, while all three models seem to be unable to capture the average returns on portfolios sorted on size and the book-to-market ratio. Finally, only C performs well in explaining portfolios sorted on momentum and generally performs best within the German stock market according to the GRS statistic. Additionally, Schrimpf et al. (2007) regard conditional versions of the CAPM and FF3 within the German stock market and Gregory et al. (2013) investigate alternative versions of FF3 and C in the UK stock market.

Fama and French (2012) examine the size, value, and momentum effects in international stock returns. Their study includes data from 23 developed markets, which they combine into the four regions North America, Japan, Asia Pacific, and Europe over the sample period 1989 to 2011. Fama and French conclude that there are value premiums in average returns in all four regions and a strong momentum effect in all regions except Japan. Moreover, they find that both value and momentum premiums are larger for small stocks (except for Japan, where there is no momentum effect in any size group). Moreover, the authors conclude that a global CAPM as well as a global FF3 and a global C (i.e., models whose factors comprise stocks from all regions) are rejected by the GRS test with global portfolios sorted on size and the book-to-market ratio and on size and momentum as testing assets. Nevertheless, when Fama and French exclude small stocks from their analysis, the global C seems to do a passable job explaining the returns on global size/book-to-market and size/momentum portfolios. However, the global models do not do well in capturing the average returns on regional portfolios sorted on size and the book-to-market ratio and on size and momentum. This finding suggests that pricing mechanisms are not sufficiently integrated across markets. When Fama and French use local models (i.e., models whose factors include stocks from only one region) to explain portfolios of stocks from the same region, the local C typically performs as well as or better than the local FF3 or the local CAPM. However, all local models seem to have low explanatory power for the size–momentum portfolios of Europe and Asia Pacific.

Although the literature on the cross-sectional variation of expected asset returns focuses on (portfolios of) stocks, some studies examine the expected returns on other asset classes. For instance, Asness et al. (1997) provide evidence for value and momentum in country equity indices, as do Bhojraj and Swaminathan (2006) for momentum and long-term reversal. Moreover, Kho (1996) and LeBaron (1999) find momentum (or, more specifically, technical trading rule profits in general) in currencies, while Erb and Harvey (2006), Miffre and Rallis (2007), and Gorton et al. (2013) show that a momentum effect is apparent in commodities. Basu and Miffre (2013) focus on the relation between hedging pressure (based on the open interest of hedgers or speculators) and expected returns on commodity futures.

Asness et al. (2013) provide an extensive analysis of value and momentum return premiums "everywhere," that is, among the eight asset classes US stocks, UK stocks, European stocks, Japanese stocks, country index futures, currencies, fixed income government bonds, and commodities. Asness et al. show that value and momentum are apparent in every asset class. Moreover, they provide evidence that the returns on value and momentum portfolios are strongly correlated across asset classes. Specifically, returns on value strategies covary positively across markets. The same applies to returns on momentum strategies. However, value and momentum returns comove negatively within and across asset classes. Motivated by this comovement across asset classes, Asness et al. propose a global three-factor model to explain the dispersion in average returns globally across asset classes, as well as locally within an asset class. Specifically, their model comprises a global market index (proxied by the MSCI World Index) as well as a zero-cost value portfolio and a zero-cost momentum portfolio, both consisting of assets from all eight classes. Asness et al. test this model over 48 high-, middle-, and low-value and momentum portfolios across all eight asset classes and show that the model performs much better than a global CAPM (although both models are rejected by the GRS test). Additionally, they show that their global three-factor model performs almost as well as FF3 in explaining the returns on the SBM25 and SM25 portfolios proposed by Fama and French (1993), which consist solely of US stocks.

2.2.3. Research questions on expected returns in the cross section of European stocks

To sum up Secs. 2.2.1 and 2.2.2, there seems to be much variation in expected returns across assets that cannot be explained by the CAPM. It looks as if multiple factors are required to capture anomalies such as size, value, and momentum. The ICAPM may be a good story to link the observation of time-varying expected returns (Sec. 2.1) to the requirement of multiple factors in the cross section. However, MSC's empirical evidence suggests that many multifactor models, including ones that are explicitly justified as empirical applications of the ICAPM by their authors, are in effect not consistent with the ICAPM, despite their ability to capture CAPM anomalies. While a variety of papers test whether the CAPM, FF3, and C also hold in markets outside the US (e.g., Fama and French, 2012) and while few non-US studies investigate FF5 (e.g., Ziegler et al., 2007), many other multifactor models have not yet been tested using data from international stock markets and other asset classes. Hence, it is unclear whether these models are valid OOS. Moreover, none of the multifactor models presented earlier has been investigated on its consistency with the ICAPM OOS, that is, whether it is consistent with the ICAPM using international stock returns or returns on other asset classes.

The second and third research questions of this thesis seek to fill this research gap to some extent. In particular, the two research questions conduct a European investigation of the eight multifactor asset pricing models that were previously tested by MSC using US data in an attempt to assess their OOS validity. First, I examine whether these models are able to explain the cross section of average returns on European stocks. Second, I test whether they are consistent with the ICAPM within the European stock market. To be specific, the second research question of this thesis is as follows.

RESEARCH QUESTION 2: Are the CAPM and the multifactor models of Fama and French (1993), Carhart (1997), Pástor and Stambaugh (2003), Campbell and Vuolteenaho (2004), Hahn and Lee (2006), Petkova (2006), and Koijen et al. (2010) able to describe the cross section of expected returns on European stocks?

Similarly, the third research question is as follows.

RESEARCH QUESTION 3: Are the multifactor models of Fama and French (1993), Carhart (1997), Pástor and Stambaugh (2003), Campbell and Vuolteenaho (2004), Hahn and Lee (2006), Petkova (2006), and Koijen et al. (2010) consistent with the ICAPM within the European stock market?

I contribute to the literature by providing an OOS test of MSC's empirical results on European data. Similar results to those of MSC would strengthen their conclusion that most multifactor models are inconsistent with the ICAPM, even though these models seem to be able to reasonably describe the cross section of average stock returns. Differing results would suggest that MSC's conclusions are not robust across different stock markets but, instead, specific to their US sample. Such results would indicate that the good explanatory power of the multifactor models under investigation might be the result of data-snooping activities, according to Lo and MacKinlay (1990). Additionally, differing results would suggest that European decision makers and regulation authorities should use a different asset pricing model for applications such as estimating the cost of equity capital (CE) than the results of MSC suggest.

2.3. Applications

Cochrane's (2011) Presidential Address to the American Finance Association emphasizes that time-varying expected returns, CAPM anomalies such as value and momentum, and multifactor asset pricing models change many practical applications. Among these applications are portfolio theory, performance evaluation, as well as corporate finance, accounting, and regulation. I first outline some aspects that affect portfolio theory and performance evaluation. Then, I emphasize an important application of asset pricing models within corporate finance, accounting, and regulation: the estimation of the CE. Afterward, I formulate the last two research questions of this thesis, which are concerned with estimating the CE for European industries using various multifactor asset pricing models.

2.3.1. Portfolio theory

Many papers are concerned with how investors should form their portfolios of assets and which trading strategies they should follow to hedge themselves against or profit from time-varying expected returns (e.g., Merton, 1971; Brennan et al., 1997; Campbell and Viceira, 1999; Barberis, 2000; Pástor, 2000; Moskowitz et al., 2012; Johannes et al., 2014). Despite this variety of papers, Cochrane (2011) emphasizes that the *average* investor must still hold the market portfolio, that is, every asset in the economy weighted by its market value, which now is multifactor efficient but not mean–variance efficient (see also Fama, 1996, and Sec. 2.2.1.2). For instance, not every investor can profit from value and momentum by buying value stocks and past winners and selling growth stocks and past losers. Instead, if there is a risk-based story behind value and momentum, there must be investors who *sell* value and past winners because these stocks are *too risky*. Similarly, not everyone can time the market. Cochrane suggests that a portfolio theory that is consistent with this theorem may be built on differences between market participants. For instance, based on Fama and French's (1996) story on value, Cochrane argues that *tech nerds* should short growth stocks, since their human capital probably correlates with these stocks, while the human capital of *steelworkers* probably covaries with value stocks, so that they sell value and buy growth. In a recent practical article, Asness and Liew (2014) emphasize, however, that they of no one who offers any systematic opposite product, such as a fund that is short in value and long in growth, a fact that speaks against a purely risk-based explanation of these effects. Overall, Cochrane states that refraining from traditional mean–variance optimization and, instead, employing multifactor optimizers, while accounting for differences among people and their hedging needs, is the great challenge for both academics and practitioners.[12]

2.3.2. Performance evaluation

From the 1970s perspective, market beta is the sole characteristic that should influence an asset's expected return. Bearing this in mind, portfolio managers should search for assets that exhibit a significant Jensen's alpha, that is, they should run time-series regressions for each potential asset i,

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_i\left(R_{m,t} - R_{f,t}\right) + e_{i,t}, \tag{2.39}$$

and buy those assets that exhibit an intercept, α_i, that is significantly positive (Jensen, 1968). Given the hypothesis that the CAPM is the true asset pricing model, a positive α_i means that asset i produces an average return that is too high relative to its risk. Likewise, regression (2.39) can be used to measure the performance of an entire portfolio or fund by using the portfolio's return. A positive intercept indicates that the portfolio manager is doing a good job, since the manager's portfolio produces higher returns than implied by the CAPM.[13]

Regression (2.39) can be replaced by multifactor models to account for CAPM anomalies such as value and momentum (e.g., Carhart, 1997; Kosowski et al., 2006; Fama and French, 2010). Given the large variety of multifactor models proposed in the finance literature and the absence of a consensus on which model is best, it is, however, unclear which of them a portfolio manager should use (e.g., Daniel and Titman, 2012). Beyond that, the focus on alpha to evaluate portfolios does, in effect, contradict Fama's (1970a, 1991) market efficiency hypothesis. In particular, under the null hypothesis that security prices fully reflect all available information, a portfolio's alpha actually reflects systematic risks that are not captured by market beta and by the slope coefficients on other factors included in the regression to estimate the alpha. This is emphasized by Cochrane (2011, p. 1087): "There is no 'alpha.' There is just beta you understand and beta you do not understand, and beta you are positioned to buy versus beta you are already exposed to and should sell."

Hence, the concept of performance evaluation seems to evolve from its earlier focus on "chasing alpha" to evaluating which systematic risks with respect to betas on multiple factors a portfolio is and should be exposed to, considering its clients' risk preferences.[14]

---

[12] In a recent paper, Cochrane (2014) shows that one can apply mean–variance portfolio theory despite time-varying expected returns if one considers long-run payoff streams such as dividends following a stock purchase, instead of one-period returns.

[13] Treynor (1965) proposes a similar concept.

[14] Aragon and Ferson (2006) provide an extensive review on performance evaluation.

2.3.3. Cost of equity capital

Time-varying expected returns and multiple factors in the cross section of expected returns change the way financial managers should estimate the CE. As pronounced by Cochrane (2011, p. 1087),

> The first slide in a capital budgeting lecture looks something like […]
>
> $$Value\ of\ investment = \frac{Expected\ payout}{R_f + \beta_m\left[E(R_m) - R_f\right]}, \tag{2.40}$$
>
> with a 6% market premium. All of which, we now know, is completely wrong. The market premium is not always 6%, but varies over time by as much as its mean. […] Expected returns do not line up with CAPM betas, but rather with multifactor betas to the extent we understand them at all. And since expected returns change over time, the discount rate is different for cashflows at different horizons.

Nevertheless, some authors still argue in favor of using the CAPM for CE estimation, despite the model seeming to fail in describing the cross section of returns. For instance, Da et al. (2012) argue that a firm's value consists of both the net present value of current projects and the value of associated real options. For this reason, stock returns (which include returns on real options) may not be explainable by the CAPM, while, at the same time, project returns (which do not include returns on real options) may follow the CAPM.

On the other hand, a few papers propose alternatives to Eq. (2.40). Most notably, Ang and Liu (2004) develop a model that is based on a conditional CAPM and that considers time-varying risk-free rates, risk premiums, and market betas. The model results in a term structure of discount rates that assigns different CEs to expected cash flows at different horizons. The authors show that the term structures of growth, neutral, and value stocks are all upward sloping at the end of December 2000. The variables that are used in their framework to estimate the conditional equity premium and betas do, however, affect the shape of the estimated term structure and hence the capital budgeting decision. Consequently, uncertainty with regard to the choice of variables (see Sec. 2.1.1) is a potential problem in applying Ang and Liu's model. In related work, Callen and Lyle (2014) model and estimate a term structure of implied CEs using synthetic futures prices derived from option contracts. Moreover, many corporate finance textbooks by now advocate the use of multifactor asset pricing models, such as FF3, to estimate CEs (e.g., Pratt and Grabowski, 2010). Regarding the multitude of multifactor models that are proposed to describe the cross section of expected returns, it is, however, completely unclear which model a financial manager should use (as unclear as the choice of model for performance evaluation pointed out in Sec. 2.3.2). For instance, regarding empirical tests of (multi-)factor models, Daniel and Titman (2012, p. 105) state

> The first concern is that these results present a conundrum for anyone attempting to use the models. Which, if any, of these dozen or so models is the correct one to use in determining cost of capital for an individual firm? The results in these papers offer no answer to this question, as each of the proposed models appears to

"work" reasonably well, in that the corresponding empirical test fails to reject the model.

Beyond that, some researchers (e.g., Cochrane, 2011) do not think the answer to practical problems such as project valuation lies in multifactor models or in dynamic present value models that incorporate time variation in expected returns—at least not yet. Instead, Cochrane advocates taking one step backward and using simpler approaches such as discounting with average returns of similar securities, as well as using "comparables" or multiples, that is, regarding the prices of known assets, such as listed firms with similar characteristics. Cochrane's hypothesis is based on FF97, who examine the CEs for 48 US industries that are obtained from the CAPM and FF3 over the period 1963 to 1994. FF97 state that the CE assessment involves at least three problems:

(1) There is no common opinion on whether the CAPM, FF3, or any other multifactor model can be called the best asset pricing model when one takes into account both their empirical performances in explaining the cross section of stock returns and their theoretical foundations. The authors state, however, that the choice of model is important, since the CE estimates of these two models typically differ by more than two percentage points per year.

(2) The estimates of risk loadings within both the CAPM and FF3 are imprecise. In particular, industries' market betas as well as their slope coefficients on SMB and HML vary strongly through time. Consequently, risk loading estimates from full sample regressions (that ignore time variation) are no more precise than estimates from regressions over short (more recent) sample periods (which are naturally imprecise, since they rely on few data points). Beyond that, the variation of individual firms' or individual projects' risk loadings will be even higher, since they do not profit from the diversification effects of portfolios.

(3) The estimates of factor risk premiums within the CAPM and FF3 are imprecise as well. For instance, FF97 emphasize that the annualized equity premium, measured as the excess return on the CRSP value-weighted market portfolio over the period 1963 to 1994, is 5.16%, with a standard error of 2.71%. Consequently, the actual equity premium might be anywhere between zero and 10%, according to the traditional rule of thumb of plus or minus two standard errors.

Overall, FF97 conclude that CE estimates "are distressingly imprecise" (p. 178) and that project valuation therefore "is beset with massive uncertainty" (p. 179). Gregory and Michou (2009, hereafter GM09) investigate whether the FF97 results apply to 35 UK industries over the sample period 1975 to 2005. The authors emphasize that CE estimation is not only important for the valuation of investments of UK firms, but also for the UK regulatory process, which uses CEs to set prices in regulated industries such as water, electricity, gas, and airports. Great uncertainty about the true CEs of UK industries therefore has potentially serious implications on UK regulatory policies. In addition to the models that FF97 examine, GM09 investigate C, as well as the four-factor model of Al-Horani et al. (2003), which augments FF3 with a research and development factor.

GM09 conclude that CEs of UK industries are estimated with large errors and the overall picture that comes out of their UK study “is every bit as bleak” (p. 701) as the picture that emerges from FF97. In particular, they state that the errors associated with estimating the slope coefficient in both the CAPM and FF3 are large and that it is difficult to decide whether the CAPM or FF3 is more qualified to estimate CEs. While forecasting the CEs three and five years forward is associated with smaller errors when GM09 use FF3, the slope coefficients on the size and value factors seem to vary considerably through time, which makes them difficult to be estimated precisely and, in contrast to the FF97 results, this variation does not seem to be due to variation in the industries’ sizes and book-to-market ratios.

Moreover, GM09 sum up that they “see nothing in [… their] results to suggest that the addition of a fourth factor, momentum, has anything significant to contribute” (p. 701), that is, C does not seem to produce CEs that are any more precise than those obtained from the CAPM and FF3. Furthermore, the authors conclude that the model of Al-Horani et al. (2003) “clearly has potential” (p. 701), since the errors of predicting CEs using this model are broadly comparable to those using FF3, while the slope coefficient of the model’s research and development factor seems to be relatively stable over time. The sample period used to estimate this model is, however, limited to 1991 to 2005 due to data availability. Finally, GM09 find that all the models investigated outperform a simple one-factor model that assumes a market beta equal to one for all industries, that is, an expected return on all industries that equals the expected return on the market portfolio. With regard to the remaining models, however, the results of GM09 provide no answer on which model financial managers and regulatory authorities should use in the UK to estimate CEs.

2.3.4. Research questions on the cost of equity capital for European industries

In summary, it appears to be beyond dispute that time-varying expected asset returns and multiple factors in the cross section of returns affect many applications of asset pricing models. It is, however, largely unclear *how* applications should change to cope with these empirical discoveries. For instance, it looks completely unclear which, if any, (multifactor) model should replace the CAPM for cost of equity (CE) estimation. In particular, while the studies of FF97 and GM09 test whether the CAPM, FF3, and C are qualified for CE estimation, it seems that none of the multifactor models presented earlier in this thesis that consider liquidity risk (Sec. 2.2.1.5) or that are explicitly justified as empirical applications of the ICAPM by their authors (Sec. 2.2.1.6) has yet been examined for their ability to estimate accurate CEs.

The last two research questions of this thesis seek to fill this research gap to some extent, focusing on the estimation of CEs. In particular, I first investigate the abilities of the CAPM, FF3, and C to precisely estimate CEs for European industries in an attempt to assess whether the results of FF97 and GM09 are valid OOS. Second, I examine whether multifactor models that consider liquidity risk or that are explicitly justified as empirical applications of the ICAPM by their authors provide more accurate CE estimates than the CAPM, FF3, or C. For this purpose, I investigate all eight multifactor models that are considered in the second and third research questions of this thesis—that is, FF5, PS, CV, HL, P, and KLVN, in addition to the CAPM, FF3, and C—on their qualifications to produce precise CEs. To be specific, the fourth research question is as follows.

RESEARCH QUESTION 4: Do the CAPM and the multifactor models of Fama and French (1993), Carhart (1997), Pástor and Stambaugh (2003), Campbell and Vuolteenaho (2004), Hahn and Lee (2006), Petkova (2006), and Koijen et al. (2010) provide precise estimates of the costs of equity capital for European industries?

Beyond that, I examine whether the choice of (multi-)factor model and estimation technique is at all important for estimating CEs for European industries. For this purpose, I formulate the last research question of this thesis as follows.

RESEARCH QUESTION 5: Do CE estimates for European industries obtained from different factor models—the CAPM and the multifactor models of Fama and French (1993), Carhart (1997), Pástor and Stambaugh (2003), Campbell and Vuolteenaho (2004), Hahn and Lee (2006), Petkova (2006), and Koijen et al. (2010)—and different estimation techniques differ from each other?

Both research questions seek to support European decision makers and regulatory authorities in deciding which asset pricing model to use to estimate CEs for European industries.

3. Data and Variables

3.1. Data and Variables for Expected Returns on Commodity Futures in the Time Series

The sample period is from January 1972 to June 2010, with a monthly sample frequency. All prices and returns are denominated in US dollars.[15]

3.1.1. Response variable

I study the predictability of one variable: the return on a portfolio that consists of several commodity futures.

Return on commodity futures, CM: I employ monthly returns on an equal-weighted portfolio of 27 commodity futures that is constructed by Asness et al. (2013). The portfolio covers aluminum, copper, nickel, zinc, lead, tin, Brent crude oil, gas oil, live cattle, feeder cattle, lean hogs, corn, soybeans, soy meal, soy oil, wheat, WTI crude, RBOB gasoline, heating oil, natural gas, gold, silver, cotton, coffee, cocoa, sugar, and platinum. The futures returns are calculated by computing the daily excess return of the most liquid futures contract every day (typically the nearest- or next nearest-to-delivery contract). The daily returns are then compounded to a total return index and the monthly returns are computed from this index. The returns do not include the return on collateral associated with the futures contract. Thus, these returns are comparable to returns in excess of the risk-free interest rate. This is important to note because we are attempting to forecast the reward for risk, not the interest rate. The data are obtained from Tobias J. Moskowitz's website.[16]

3.1.2. Potential predictor variables

The core of Research Question 1 is a test of the null hypothesis that returns on commodity futures are unpredictable against the alternative hypothesis that the expected returns depend on factors such as price levels and past price movements, economic conditions, and investor sentiment and consequently vary through time. Accordingly, my approach is not to test a specific theory of commodity futures returns that represents this alternative hypothesis and, at the same time, predetermines the set of potential predictor variables but, rather, to choose the candidate predictors myself. While this approach examines variables that have not yet been suggested by any theory, the drawback of this approach is the selection of potential predictors that is, to some extent, arbitrary.

I employ a total of 32 variables that reflect price levels and past price movements, economic conditions, and investor sentiment (see Table 4).

[15] This chapter is based on Sec. 2 of Lutzenberger (2014a).
[16] See http://faculty.chicagobooth.edu/tobias.moskowitz/research/data.html.

Table 4. Potential predictors of commodity futures returns.

Variable	Description	Examples of studies	Proxy	Data source	Sample period
Panel A: Stock market					
DY	Dividend–price ratio	Stocks: Campbell and Shiller (1988a); Fama and French (1988b)	Log of S&P 500 dividend–price ratio (Welch and Goyal, 2008)	Amit Goyal[17]	1/1972–6/2010
E/P	Earnings–price ratio	Stocks: Campbell and Shiller (1988b)	Log of S&P 500 earnings–price ratio (Welch and Goyal, 2008)	Amit Goyal	1/1972–6/2010
D/E	Dividend–payout ratio	Stocks: Lamont (1998)	Log of S&P 500 dividend–earnings ratio (Welch and Goyal, 2008)	Amit Goyal	1/1972–6/2010
SVAR	Stock variance	Stocks: Guo (2006a)	Sum of squared daily S&P 500 returns (Welch and Goyal, 2008)	Amit Goyal	1/1972–6/2010
B/M	Book-to-market ratio	Stocks: Kothari and Shanken (1997); Pontiff and Schall (1998)	Dow Jones Industrial Average book–market ratio (Welch and Goyal, 2008)	Amit Goyal	1/1972–6/2010
CSP	Cross-sectional premium	Stocks: Polk et al. (2006)	-	Amit Goyal	1/1972–2/2002
CRMRF	Cumulative equity premium	-	Five-year cumulative sum of Fama and French’s (1993) US market excess return	Kenneth R. French	1/1972–6/2010
CL	Stock liquidity	Stocks: Maio and Santa-Clara (2012)	Five-year cumulative sum of Pástor and Stambaugh’s (2003) non-traded liquidity factor	Lubos Pástor	1/1972–6/2010
SENT	Investor sentiment	Stocks: Baker and Wurgler (2006; 2007)	Stock market sentiment index of Baker and Wurgler (2006, 2007)	Jeffrey Wurgler	1/1972–6/2010
Panel B: Bond market					
RF	Treasury bills	Stocks: Campbell (1987)	Three-month US Treasury bill rate (secondary market) (Welch and Goyal, 2008)	Amit Goyal	1/1972–6/2010
LTY	Long-term yield	Stocks: Welch and Goyal (2008)	Yield on long-term US government bonds (Welch and Goyal, 2008)	Amit Goyal	1/1972–6/2010
LTR	Long-term return	Stocks: Welch and Goyal (2008)	Return on long-term US government bonds (Welch and Goyal, 2008)	Amit Goyal	1/1972–6/2010
TERM	Term structure spread	Stocks: Campbell (1987); Fama and French (1989)	Long-term yield minus US Treasury bill rate (Welch and Goyal, 2008)	Amit Goyal	1/1972–6/2010
DEF	Default yield spread	Stocks: Fama and French (1989); Keim and Stambaugh (1986)	Yield on BAA- minus yield on AAA-rated corporate bonds (Welch and Goyal, 2008)	Amit Goyal	1/1972–6/2010
DFR	Default return spread	Stocks: Welch and Goyal (2008)	Return on long-term corporate bonds minus return on long-term US government bonds (Welch and Goyal, 2008)	Amit Goyal	1/1972–6/2010
CP	Cochrane–Piazzesi factor	Bonds and stocks: Cochrane and Piazzesi (2005) Stocks: Maio and Santa-Clara (2012)	-	John H. Cochrane	1/1972–12/2003

[17] See http://www.hec.unil.ch/agoyal/.

Table 4 (continued)

Variable	Description	Examples of studies	Proxy	Data source	Sample period
Panel C: Macroeconomics					
INFL	Inflation	Stocks: Fama (1981)	US Consumer Price Index inflation (all urban consumers) lagged by one month (Welch and Goyal, 2008)	Amit Goyal	1/1972–6/2010
I/K	Investment-to-capital ratio	Stocks: Cochrane (1991)	(quarterly data are linearly interpolated to obtain monthly data)	Amit Goyal	1/1972–6/2010
CAY	Consumption–wealth ratio	Stocks: Lettau and Ludvigson (2001a)	(quarterly data are linearly interpolated to obtain monthly data)	Amit Goyal	1/1972–6/2010
IP	Industrial production	Commodities (volatility): Prokopczuk and Symeonidis (2013)	Five-year log growth of US industrial production	Federal Reserve Bank of St. Louis	1/1972–6/2010
M2	M2 money stock	Commodities (volatility): Prokopczuk and Symeonidis (2013)	Three-year log growth of the US M2 money stock	Federal Reserve Bank of St. Louis	1/1972–6/2010
GDP	Gross domestic product (GDP)	Stocks: Rangvid (2006)	Three-year log growth of the US GDP (quarterly data are linearly interpolated to obtain monthly data)	Federal Reserve Bank of St. Louis	1/1972–6/2010
USD	Return on US dollar	Commodities (volatility): Prokopczuk and Symeonidis (2013)	Five-year log return on a trade-weighted US dollar index against major currencies	Federal Reserve Bank of St. Louis	1/1978–6/2010
CVAL	Value everywhere factor	-	Five-year log excess return on the value everywhere factor of Asness et al. (2013)	Tobias J. Moskowitz	12/1976–6/2010
CMOM	Momentum everywhere factor	-	Five-year log excess return on the momentum everywhere factor of Asness et al. (2013)	Tobias J. Moskowitz	12/1976–6/2010
Panel D: Commodity market					
CVAR	Commodity variance	-	Sum of squared daily CRB BLS spot index returns	Datastream	1/1972–6/2010
CCM_spot	Commodity spot return	-	Five-year log return on CRB BLS spot index	Datastream	1/1972–6/2010
CCM	Commodity futures return	-	Five-year log return on the commodity futures portfolio of Asness et al. (2013)	Tobias J. Moskowitz	12/1976–6/2010
C12CM_spot	Commodity spot momentum	-	12-month log return on CRB BLS spot index (most recent month's return is skipped)	Datastream	1/1972–6/2010
C12CM	Commodity futures momentum	Several asset classes: Moskowitz et al. (2012)	12-month log return on the commodity futures portfolio of Asness et al. (2013) (most recent month's return is skipped)	Tobias J. Moskowitz	12/1972–6/2010
CVAL_CM	Commodity value factor	-	Five-year log excess return on the commodity value factor of Asness et al. (2013)	Tobias J. Moskowitz	12/1976–6/2010
CMOM_CM	Commodity momentum factor	-	Five-year log excess return on the commodity momentum factor of Asness et al. (2013)	Tobias J. Moskowitz	12/1976–6/2010

The variables in Table 4 are classified into four groups: stock characteristics (Panel A), bond characteristics (Panel B), macroeconomic characteristics (Panel C), and commodity characteristics (Panel D). A large number of these variables are relatively common choices in the literature that studies the predictability of stock and bond returns (see Sec. 2.1). I am interested in whether these variables also have forecasting power over commodity futures returns and, therefore, I include them in the set of candidate predictors. Because space is limited, however, I refer to the studies mentioned in Table 4 for a description of the motivation behind these variables and their construction. Some variables are less standard choices in the predictability of returns literature or have not yet been considered and are, therefore, presented in more detail herein.

The first three potential predictive variables, which I present at length, are stock market characteristics.

Equity premium (five-year cumulative sum), CRMRF: Cumulative sum of the equity premium, which is the total return on the stock market in excess of the risk-free rate over the last 60 months. I obtain monthly data for the US equity premium, since it is employed by Fama and French (1993), from Kenneth R. French's website, which includes all NYSE, AMEX, and NASDAQ firms. The intention behind the cumulative sum is to obtain a slow-moving predictive variable that corresponds to the equity premium. I choose the last 60 months rather than the total cumulative sum (or index level) because the variable constructed in this way is stationary, whereas the total cumulative sum is close to being non-stationary (the autocorrelation coefficient is around one). This approach is proposed by Maio and Santa-Clara (2012, hereafter MSC), who construct a predictive (state) variable that is associated with the momentum factor of Carhart (1997), as well as the liquidity factor mentioned in the following. I also apply this approach to several other return or growth rate series for which I want to obtain associated slow-moving predictive variables.

Liquidity factor (five-year cumulative sum), CL: Cumulative sum of Pástor and Stambaugh's (2003) non-traded liquidity factor over the last 60 months, which represents innovations in aggregate stock market liquidity (Pástor and Stambaugh 2003, equation (8)), obtained from Lubos Pástor's website.[18] This variable is employed by MSC.

Investor sentiment, SENT: The stock market sentiment index of Baker and Wurgler (2006, 2007) is based on the first principal component of six US sentiment proxies: the NYSE turnover, the dividend premium, the closed-end fund discount, the number of initial public offerings and their first-day returns, and the equity share in new issues. The monthly data are obtained from Jeffrey Wurgler's website[19] and are described by Baker and Wurgler (2007). I choose the series where each of the proxies has first been orthogonalized with respect to a set of macroeconomic conditions.

I then highlight a potential predictor that is a bond market characteristic.

---

18 See http://faculty.chicagobooth.edu/lubos.pastor/research/.

19 See http://people.stern.nyu.edu/jwurgler/.

Cochrane–Piazzesi factor, CP: The factor proposed by Cochrane and Piazzesi (2005) is the fitted value from a regression of an average of excess bond returns on forward rates and is related to bond risk premiums. I obtain the necessary data to construct *CP*, which cover the period January 1972 to December 2003, from John H. Cochrane's website.[20]

The next set of independent variables, which I explicitly outline below, can be categorized as primarily macroeconomic factors.

Industrial production growth (five-year), IP: Five-year log growth of US industrial production, for which the data are obtained from the Federal Reserve Bank of St. Louis,[21] as are the data for the following three variables.

M2 money stock growth (three-year), M2: Three-year log growth of the US M2 money stock.

GDP growth (three-year), GDP: Three-year log growth of the US GDP. I linearly interpolate the quarterly GDP data to obtain a monthly series.

Return on US dollar (five-year), USD: Five-year log return on a trade-weighted US dollar index against major currencies. The series covers January 1978 to June 2010.

Return on Value Everywhere (five-year), CVAL: Cumulative sum of the log excess return on the value everywhere factor of Asness et al. (2013) over the last 60 months. The factor comprises eight asset classes (US equities, UK equities, continental European equities, Japanese equities, global equity indices, currencies, fixed income, and commodities). The variable seeks to represent the cross-sectional value return premium across these eight asset classes. I obtain the factor data from Tobias J. Moskowitz's website. The series thus constructed covers December 1976 to June 2010.

Return on Momentum Everywhere (five-year), CMOM: Cumulative sum of the log excess return on the momentum everywhere factor of Asness et al. (2013) over the last 60 months. The factor comprises the same eight asset classes as the value everywhere factor. It represents the cross-sectional momentum return premium across these eight asset classes. The data are from Tobias J. Moskowitz's website. The resulting series covers December 1976 to June 2010.

Finally, I construct the following various factors from commodity market data that potentially show predictive power over commodity futures returns:

Commodity variance, CVAR: In a manner analogous to that for *SVAR*, I compute the volatility of the aggregate commodity spot market as the sum of squared daily returns on the CRB BLS spot index. The price index data are obtained from Thomson Reuters Datastream.

Return on CRB BLS spot index (five-year), CCM_spot: Cumulative sum of the monthly log return on the CRB BLS spot index over the last 60 months. The intuition behind this variable is to capture potential time-series value in commodities (for more detail on the cross-sectional value effect in commodities, see Asness et al., 2013).

[20] See http://faculty.chicagobooth.edu/john.cochrane/research/Data_and_Programs/Bond_Risk_Premia/.
[21] See http://research.stlouisfed.org/.

Return on commodity futures (five-year), CCM: Cumulative sum of *CM* over the last 60 months. This variable also seeks to capture potential time-series value. The series computed thus covers December 1976 to June 2010.

Return on CRB BLS spot index (12-month), C12CM_spot: Cumulative sum of the monthly log return on the CRB BLS spot index over the last 12 months, where the most recent month's return is skipped. This measure is the common measure to capture momentum (Jegadeesh and Titman, 1993; Asness et al., 2013). Significant time-series momentum in commodity futures is found by Moskowitz et al. (2012).

Return on commodity futures (12-month), C12CM: Cumulative sum of *CM* over the last 12 months, where the most recent month's return is skipped. This variable is also employed to capture time-series momentum. The resulting series covers December 1972 to June 2010.

Return on Commodities Value (five-year), CVAL_CM: Cumulative sum of the log excess return on the commodities value factor of Asness et al. (2013) over the last 60 months. The variable represents the cross-sectional value return premium in commodities. The data are from Tobias J. Moskowitz's website, as they are for the following variable. The series constructed this way covers December 1976 to June 2010.

Return on Commodities Momentum (five-year), CMOM_CM: Cumulative sum of the log excess return on the commodities momentum factor of Asness et al. (2013) over the last 60 months. The factor represents the cross-sectional momentum return premium in commodities. The series covers December 1976 to June 2010.

3.1.3. Summary statistics

Table 5 shows summary statistics for the response and predictor variables. Observe that the first-order autoregressive coefficients of most predictors are above 0.9, indicating that most predictors are highly persistent. Some variables are correlated with others to some extent (available upon request). Correlation coefficients above 0.85 are shown by *DY* and *E/P* (0.91), *DY* and *B/M* (0.91), *E/P* and *B/M* (0.91), *B/M* and *GDP* (0.87), and *RF* and *LTY* (0.86).

Table 5. Summary statistics for response and predictor variables.
This table reports the mean, standard deviation (Std.) and first-order autocorrelation coefficient (AC_1) of the response and predictor variables employed in this study. The sample period for the majority of variables is January 1972 to June 2010. Some variables are only available for a shorter sample period. The data series as well as the sources are described in Secs. 3.1.1 and 3.1.2.

Variable	Mean	Std.	AC_1
CM	0.0044	0.0433	0.133
DY	-3.5801	0.4508	0.995
E/P	-2.8186	0.5118	0.990
D/E	-0.7615	0.3426	0.983
SVAR	0.0025	0.0051	0.464
B/M	0.5117	0.3008	0.995
CSP	-0.0013	0.0010	0.947
CRMRF	0.2043	0.3220	0.977
CL	-0.0672	0.5828	0.991
SENT	-0.0260	0.9205	0.986
RF	0.0561	0.0312	0.987
LTY	0.0761	0.0245	0.990
LTR	0.0074	0.0311	0.039
TERM	0.0200	0.0153	0.947
DEF	0.0111	0.0048	0.963
DFR	-0.0001	0.0141	-0.012
CP	0.0113	0.0242	0.741
INFL	0.0036	0.0038	0.618
I/K	0.0362	0.0035	0.997
CAY	0.0031	0.0224	0.995
IP	0.1209	0.0800	0.992
M2	0.1999	0.0751	0.997
GDP	0.2029	0.0702	0.998
USD	-0.0424	0.1859	0.988
CVAL	0.2003	0.1487	0.972
CMOM	0.3530	0.2171	0.986
CCMspot	0.1636	0.2694	0.991
C12CMspot	0.0316	0.1297	0.957
CVAR	0.0004	0.0006	0.504
CCM	0.1774	0.3906	0.989
C12CM	0.0476	0.1926	0.957
CVAL_CM	0.1558	0.5979	0.988
CMOM_CM	0.5514	0.4122	0.978

3.2. Data for Expected Returns in the Cross Section of European Stocks

I obtain from Datastream end-of-month return index (Datastream variable RI), market value (Datastream variable MV), and price-to-book value (Datastream variable PTBV) data, as well as data for the daily return index, price (Datastream variable P) and turnover by volume (Datastream variable VO) for all stocks of 16 European countries (Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, and the United Kingdom) that are categorized as primary quote, major security, and equity and for which Datastream offers adjusted prices.[22] The data cover the period from November 1989 to December 2011. This rather short sample period ensures broad coverage of stocks. I include dead and suspended companies to avoid

[22] This chapter is based on Sec. 4 of Lutzenberger (2014b).

survivorship bias and the data are denominated in euros. The data include utilities and financials, as for Fama and French (2012). The overall sample consists of 19,226 stocks. The daily data are required only for the estimation of Pástor and Stambaugh's (2003) model. The 16 European countries I choose are used as a representation of Europe, for example, by Bauer et al. (2010), Fama and French (2012), and Wallmeier and Tauscher (2014). On average, they account for 30% of global market capitalization (Fama and French, 2012).

In the monthly data set, I exclude all stock months (i.e., data points of an individual stock) for which I do not have both market value and return index data. Second, I follow Guo and Savickas (2008) and impose additional filters on the Datastream data to remove potential coding errors.[23] These authors obtain daily Datastream data for all G7 countries and obtain, for the US, essentially the same results from filtered Datastream data as from CRSP data. I apply one of their five filters for daily data to my monthly data. In particular, if the return index of a stock is below three in a month, I exclude this stock month from the analysis.

In the daily data set, I include only stock days (i.e., the data points of an individual stock) for which I have both the return index and the price and turnover by volume data. I then impose all five filters suggested by Guo and Savickas (2008). Specifically, if the return index of a stock is below three in a day, I exclude this stock day from the analysis. Second, if the simple daily return on a stock is greater than 300%, that stock day is excluded. Third, I exclude stock days in which the absolute value of changes in price is more than 50% in one day.[24] Fourth, if a stock's price falls by more than 90% in a day and has increased by more than 200% within the previous 20 days, all stock days between the two dates are excluded. Fifth, if a stock's price increases by more than 100% in a day and has decreased by more than 200% within the previous 20 days, I exclude all stock days between the two dates from the analysis.

The test assets are 25 portfolios formed on size and book-to-market ratio (SBM25) and 25 portfolios formed on size and momentum (SM25). I largely follow Fama and French (2012) in their construction. To form SBM25, at the end of June of each year t I sort the stocks into five size groups and five book-to-market groups on market value for the end of June of year t and on the book-to-market ratio for the end of December of year $t-1$, where the book-to-market ratio is the reciprocal of the price-to-book value. In doing so, I include only stock months for which I have price-to-book value data for the end of last December. Moreover, to remove potential Datastream coding errors, I require a price-to-book value to be between 3% and 50. Otherwise, a stock is not considered in the portfolio formation at t. To avoid the domination of sorts by tiny stocks, Fama and French (2012) use the third, seventh, 13th, and 25th percentiles of the aggregate market capitalization as size breakpoints. As for these authors, 75% of the stocks that are included in the portfolios formed at t have a market value lower than the first size breakpoint, 87% have a market value lower than the second breakpoint, and so on. Moreover, the book-to-market breakpoints are the 20th, 40th, 60th, and 80th percentiles of the book-to-market ratio for the stocks that are in the top 90% of the June aggregate market value. Thus, the book-to-market breakpoints are based on large stocks, again to avoid sorts

[23] An alternative procedure for screening Datastream data is provided by Ince and Porter (2006).
[24] In this third filter, Guo and Savickas (2008) consider the capitalization instead of the price.

dominated by tiny stocks. The resulting 25 value-weighted size/book-to-market portfolios are the intersections of the independent 5x5 size and book-to-market sorts.

In the next step, at the end of each month t, I sort the stocks into five size groups and five lagged momentum return groups to form SM25. I use the same breakpoint conventions as for SBM25. However, the portfolios are formed monthly and the lagged momentum return substitutes for the book-to-market ratio. For the portfolio formation at the end of month $t-1$, the lagged momentum return is computed as a stock's cumulative simple return from months $t-12$ to $t-2$. Thus, in a portfolio for month t (which is formed at the end of month $t-1$), I include only stock months for which I have a return index value for the end of month $t-13$ and for month $t-2$. The 25 value-weighted size/momentum portfolios are the intersections of the independent 5x5 sorts. Table 6 presents the summary statistics for the portfolios.

Furthermore, each month t, I form a portfolio that includes all the stocks of my monthly data set, with each stock weighted by its market value that month. I calculate the portfolio's simple monthly return in month $t+1$ and use it as a proxy for the market return. Following MSC, I add the market return to each set of test assets. To obtain excess returns, I subtract the three-month Frankfurt Interbank Offered Rate (FIBOR) until 1998 (Bauer et al., 2010), which I obtain from Deutsche Bundesbank.[25] From 1999 onward, I use the three-month Euro Interbank Offered Rate (EURIBOR) as a proxy for the risk-free rate, obtained from Datastream. I compute the realized stock market variance, $SVAR$, as the sum of squared daily returns on the Datastream Global Equity Index for Europe.[26]

To construct the Fama–French (1993) factors SMB and HML, as well as Carhart's (1997) momentum factor UMD, I again mainly follow Fama and French (2012). In particular, at the end of each June of each year t, I sort the stocks of my monthly data set into two market value groups and three book-to-market groups. I categorize stocks in the top 90% of June of year t's aggregate market value as big stocks and those in the bottom 10% as small stocks. As book-to-market breakpoints for the big and small stocks, I use the 30th and 70th percentiles, respectively, of December of year $t-1$'s book-to-market value for the big stocks. In all other respects, the grouping procedure follows the same rules as the 5x5 sorts. Next, I form six value-weighted portfolios from the intersections of the independent 2x3 sorts: small–growth (SG), small–neutral (SN), small–value (SV), big–growth (BG), big–neutral (BN), and big–value (BV). I compute SMB as the equal-weighted average of the simple returns on the three portfolios that include the small stocks (SV, SN, and SG) minus the average of the returns on the three big stock portfolios (BV, BN, and BG):

$$SMB_t = \frac{1}{3}\left(R_{SV,t} + R_{SN,t} + R_{SG,t}\right) - \frac{1}{3}\left(R_{BV,t} + R_{BN,t} + R_{BG,t}\right). \tag{3.1}$$

[25] See http://www.bundesbank.de/.

[26] This equity index has a correlation of 0.99 with my self-constructed market index.

Table 6. Summary statistics for 25 size/book-to-market and 25 size/momentum portfolios.

This table shows the summary statistics for the 25 size/book-to-market portfolios (Panel A) and 25 size/momentum portfolios (Panel B) of the European stock market that are used as test assets. The sample period is December 1990 to December 2011. The table presents the average monthly excess returns (in percent), the monthly standard deviation of the excess returns (in percent), as well as the average number of stocks in each portfolio. I subtract the three-month FIBOR rate (until 1998) or the three-month EURIBOR rate (from 1999 onward) to obtain excess returns. The rows refer to market value (size) quintiles and the columns to book-to-market quintiles (Panel A) or momentum quintiles (Panel B).

Panel A: 25 portfolios formed on size and book to market					
	Low	2	3	4	High
	Mean excess return (monthly, %)				
Small	-0.22	0.00	0.25	0.50	0.77
2	0.02	0.30	0.33	0.65	0.82
3	0.25	0.41	0.62	0.60	0.77
4	0.21	0.54	0.60	0.64	0.66
Big	0.38	0.48	0.49	0.79	0.37
	Standard deviation of excess returns (monthly, %)				
Small	5.77	5.14	4.70	4.54	4.48
2	5.76	5.04	4.80	5.00	5.12
3	5.78	4.66	4.85	4.83	5.45
4	5.36	4.83	4.92	5.23	5.75
Big	4.86	4.59	5.06	5.45	6.75
	Average number of stocks				
Small	428	420	464	582	1352
2	92	93	96	102	141
3	53	51	50	52	57
4	36	33	36	35	37
Big	27	31	31	26	18

Panel B: 25 portfolios formed on size and momentum					
	Low	2	3	4	High
	Mean excess return (monthly, %)				
Small	-0.40	0.17	0.37	0.75	1.40
2	-0.18	0.30	0.45	0.65	1.04
3	0.04	0.33	0.59	0.74	0.88
4	0.18	0.40	0.63	0.50	0.89
Big	0.02	0.32	0.56	0.60	0.63
	Standard deviation of excess returns (monthly, %)				
Small	6.01	4.25	3.74	3.71	4.83
2	6.52	4.64	4.18	4.12	5.01
3	6.43	4.94	4.87	4.30	4.92
4	6.81	5.02	4.49	4.39	5.16
Big	7.26	5.24	4.38	4.47	5.38
	Average number of stocks				
Small	1555	623	504	469	710
2	151	115	107	107	140
3	66	60	58	59	68
4	40	42	41	42	41
Big	25	33	37	35	26

The factor HML is calculated as the equal-weighted average of the simple returns on the two portfolios consisting of the high–book-to-market stocks (SV and BV) minus the average of the returns on the two low–book-to-market portfolios (SG and BG):

$$HML_t = \frac{1}{2}\left(R_{SV,t} + R_{BV,t}\right) - \frac{1}{2}\left(R_{SG,t} + R_{BG,t}\right). \tag{3.2}$$

To compute UMD, I construct 2x3 sorts on size and lagged momentum. The construction follows that of the 25 size/momentum portfolios, but the size breakpoints are those of the 2x3 sorts described above. Moreover, the momentum breakpoints are the 30th and 70th percentiles of the lagged momentum returns of the big stocks. I obtain six value-weighted portfolios as the intersections of the independent 2x3 sorts on size and momentum: small–loser (SL), small–neutral (SN), small–winner (SW), big–loser (BL), big–neutral (BN), and big–winner (BW). The factor UMD is the equal-weighted average of the simple returns for the two portfolios with high lagged momentum (SW and BW) minus the simple returns of the two portfolios with low lagged momentum (SL and BL):

$$UMD_t = \frac{1}{2}\left(R_{SW,t} + R_{BW,t}\right) - \frac{1}{2}\left(R_{SL,t} + R_{BL,t}\right). \tag{3.3}$$

I use the daily data set to construct Pástor and Stambaugh's (2003) liquidity-related factor, L. As for MSC, I use the non-traded factor (Eq. (2.17) and Pástor and Stambaugh, 2003, Eq. (8)). I follow the authors step by step in the construction of L, except that I proxy for a stock's euro volume on a given day by multiplying its closing price that day by its turnover by volume that day (i.e., by the number of its shares traded that day), since the availability of exact euro volume data in Datastream is very limited. Furthermore, Pástor and Stambaugh exclude all stocks with prices less than US\$5 and greater than US\$1,000. I adjust those numbers to four euros and 800 euros.

Table 7 displays the summary statistics for the market excess return and for the four empirical factors. The market equity premium is positive and the t-statistic of the mean indicates that it is significantly different from zero. It is higher than that presented by Fama and French (2012) for Europe, which uses the sample period November 1990 to March 2011. The means of HML and UMD are both significant and positive. The values are comparable to the values of Fama and French (2012). However, I obtain a significantly negative mean for SMB, whereas Fama and French (2012) present an average SMB that is not significantly different from zero, although its sign is also negative. Thus, I detect a negative size effect for the European stock market within my sample. This coincides with the results of Wallmeier and Tauscher (2014), who also indicate a negative size effect for Europe. Hence, it looks as if the size effect that was significantly positive in the past (Banz, 1981) but vanished over the years in the US (e.g., Fama and French, 2012) has by now reversed in Europe, inducing a negative size effect. Finally, the mean of L is not significantly different from zero.

Table 7. Summary statistics for market and empirical factors.
This table presents the summary statistics for the value-weighted market excess return RM, the size factor SMB, the value factor HML, the momentum factor UMD, and the liquidity-related risk factor L of the European stock market. The second row contains the monthly mean (in percent), the third row presents the monthly standard deviation (in percent), and the last row shows the t-statistic of the mean, that is, the ratio of the mean to its standard error. The sample period is December 1990 to December 2011. Excess returns were obtained by subtracting the three-month German FIBOR rate (until 1998) or the three-month EURIBOR rate (from 1999 onward).

	RM	SMB	HML	UMD	L
Mean (monthly, %)	1.12	-0.43	0.66	0.93	0.00
Std. (monthly, %)	4.44	2.34	2.66	4.22	0.20
t-Mean	4.00	-2.95	3.93	3.51	-0.17

I also have to find European proxies for the state variables included in the multifactor models I investigate. As a proxy for the term structure spread, $TERM$, I use the spread between the yield on 10-year German government bonds, which I obtain from Deutsche Bundesbank (series WZ9826), and the three-month FIBOR rate (until 1998; Bauer et al., 2010), and the EURIBOR rate (from 1999 onward). For the default yield spread, DEF, I take the yield spread between Moody's Baa- and Aaa corporate bonds (Bauer et al., 2010) and choose the dividend yield of the Datastream Global Equity Index for Europe for the aggregate dividend–price ratio, DY. Moreover, for the short-term risk-free rate, RF, I take the three-month FIBOR or EURIBOR rate. The aggregate price–earnings ratio, PE, is measured according to Campbell and Vuolteenaho (2004), by calculating the log ratio of the market value of the Datastream Global Equity Index for Europe to a 10-year moving average of earnings. The earnings were thereby computed from the price–earnings ratio of the Datastream Global Equity Index for Europe. The small-stock value spread, VS, is constructed as in the appendix of Campbell and Vuolteenaho (2004). The calculation is based on the difference between the log book-to-market ratios of the small–value portfolio (SV) and small–growth portfolio (SG). I follow Cochrane and Piazzesi (2005) and Eq. (2.5) to compute the Cochrane–Piazzesi factor, CP. Therefore, I obtain the term structure of interest rates on listed German Federal securities with residual maturities of one to five years from Deutsche Bundesbank (series WZ9808, WZ9810, WZ9812, WZ9814, WZ9816).

In addition to the state variables described above, I have to obtain state variables that are associated with the empirically motivated factors HML, SMB , UMD, and L. I follow MSC in their construction. I use the cumulative sums of UMD and L to compute the state variables that are associated with those two factors, $CUMD$ and CL. For example, in the case of L, I obtain the associated state variable CL as

$$CL_t = \sum_{s=t-59}^{t} L_s \tag{3.4}$$

and the state variable associated with UMD is computed similarly. I thereby lose 59 observations, so that the sample period for the predictive regressions containing $CUMD$ and CL becomes November 1995 to December 2011. To obtain the state variable corresponding to SMB, I compute the difference between the market-to-book ratios of the three small-stock portfolios, SG, SN, and SV, and the market-to-book ratios of the three big-stock portfolios, BG, BN, and BV, in each month:

$$SMB^{*} = \frac{MB_{SG} + MB_{SN} + MB_{SV}}{3} - \frac{MB_{BG} + MB_{BN} + MB_{BV}}{3}, \tag{3.5}$$

where, for example, MB_{SG} denotes the monthly market-to-book ratio of the small–growth portfolio, SG. In doing so, I compute a portfolio's monthly market-to-book ratio using all its stocks for which I have a price-to-book value between 3% and 50 for that month. The state variable associated with HML is constructed in the same style, that is, as the difference between the monthly market-to-book ratios of the value and growth portfolios:

$$HML^{*} = \frac{MB_{SV} + MB_{BV}}{2} - \frac{MB_{SG} + MB_{BG}}{2}. \tag{3.6}$$

Table 8 presents summary statistics for the state variables.

Table 8. Summary statistics for state variables.
This table displays the summary statistics for the state variables of the European stock market. The state variables are the slope of the yield curve $TERM$, the corporate bond default spread DEF, the market dividend-to-price ratio DY, the short-term risk-free rate RF, the aggregate price–earnings ratio PE, the value spread VS, the Cochrane–Piazzesi factor CP, the size premium SMB^{*}, the value premium HML^{*}, the momentum premium $CUMD$, and the liquidity factor CL. The table shows the mean, standard deviation, minimum, maximum, and first-order autocorrelation coefficient ϕ. The sample period is December 1990 to December 2011 for all variables except CL and $CUMD$, whose sample period is November 1995 to December 2011.

Variable	Mean	Std.	Min.	Max.	ϕ
$TERM$	0.012	0.014	-0.020	0.035	0.984
DEF	0.010	0.004	0.005	0.034	0.957
DY	2.949	0.717	1.730	5.840	0.974
RF	0.040	0.023	0.006	0.099	0.997
PE	3.231	0.383	2.269	4.011	0.992
VS	1.660	0.159	1.406	2.231	0.930
CP	0.014	0.008	-0.010	0.036	0.876
SMB^{*}	-0.121	0.390	-1.528	0.550	0.931
HML^{*}	-4.213	1.138	-8.841	-2.896	0.954
$CUMD$	0.537	0.183	0.173	1.054	0.945
CL	0.003	0.015	-0.037	0.023	0.978

3.3. Data for the Cost of Equity Capital for European Industries

To examine Research Questions 4 and 5, I use the European data set described in Sec. 3.2.[27] There are only a few additional data. First, I choose August 1990 to December 2011 as the sample period. Second, I employ several Datastream Global Equity Indices for Europe, that is, two levels of this index family. In particular, I choose the first-level index as the market portfolio and the fourth-level indices as the industry portfolios. The latter divide the market into 35 sectors according to the FTSE Industry Classification Benchmark.[28] Gregory and

[27] This chapter is based on Sec. 3 of Lutzenberger (2014c).
[28] An exemplary former study that makes use of such Datastream indices is Artmann et al. (2012b), which employs German industry sectors as testing assets.

Michou (2009) also use this classification system to define their UK industries. I obtain the end-of-month return indices for both the market portfolio and the industry portfolios. I compute the excess returns by subtracting the three-month German FIBOR rate until 1998 and the three-month EURIBOR rate from 1999 onward from the simple returns on the market portfolio and the industry portfolios. Table 9 provides summary statistics for the industry returns. All state variables or factors that are included in the models investigated in Research Questions 4 and 5 are constructed as described in Sec. 3.2 (except for the excess market return, which is proxied by the excess return on the first-level index described above).

Table 9. Industry summary statistics.
This table reports for each industry the mean, median, and standard deviation of monthly returns (in percent), as well as the t-statistic of the mean monthly returns, that is, the ratio of the mean to its standard error. I compute simple returns from August 1990 to December 2011.

Industry name	Industry code	Mean (monthly, %)	Median (monthly, %)	Std. (monthly, %)	t-Mean
Oil & Gas	1	0.71	1.15	5.48	2.06
Chemicals	2	0.69	1.13	5.51	2.02
Forestry & Paper	3	0.20	0.20	6.48	0.49
Industrial Metals	4	0.54	0.58	8.24	1.05
Mining	5	1.02	1.25	8.30	1.97
Construction & Materials	6	0.31	1.23	5.78	0.87
Aerospace & Defense	7	0.55	0.84	6.38	1.39
General Industrials	8	0.43	1.10	5.73	1.20
Electronic & Electrical Equipment	9	0.63	1.04	7.59	1.33
Industrial Engineering	10	0.59	0.76	6.24	1.53
Industrial Transportation	11	0.34	1.31	5.27	1.04
Support Services	12	0.53	1.46	5.37	1.58
Automobiles & Parts	13	0.38	0.28	7.35	0.83
Beverages	14	0.63	1.35	4.47	2.25
Food Producers	15	0.56	0.76	3.85	2.32
Household Goods	16	0.50	1.34	5.12	1.56
Leisure Goods	17	0.37	0.72	6.16	0.95
Personal Goods	18	0.78	1.31	5.39	2.31
Tobacco	19	1.28	1.75	5.45	3.76
Health Care Equipment & Services	20	0.59	1.07	4.40	2.14
Pharmaceuticals & Biotechnology	21	0.70	1.02	4.13	2.71
Food & Drug Retailers	22	0.39	0.66	4.46	1.41
General Retailers	23	0.51	0.95	5.11	1.61
Media	24	0.36	0.71	6.31	0.92
Travel & Leisure	25	0.17	0.63	5.94	0.47
Fixed Line Telecommunications & Mobile Telecommunications	26	0.55	0.79	6.28	1.40
Electricity	27	0.55	0.74	4.13	2.15
Gas, Water, & Multiutilities	28	0.48	0.84	4.41	1.75
Banks	29	0.30	0.95	6.55	0.74
Non-Life Insurance	30	0.28	0.85	6.49	0.68
Life Insurance	31	0.27	0.86	7.39	0.58
Real Estate	32	0.23	0.79	4.64	0.78
General Financial	33	0.34	0.68	4.94	1.11
Software & Computer Services	34	0.82	1.13	8.70	1.50
Technology Hardware & Equipment	35	0.63	0.48	9.88	1.02

4. Empirical Methodology

4.1. Methodology for Expected Returns on Commodity Futures in the Time Series

I conduct both an in-sample (IS) analysis and an out-of-sample (OOS) analysis to investigate whether aggregate returns on commodity futures are predictable.[29]

4.1.1. In-sample prediction

I begin with IS single long-horizon predictive regressions, which are the common approach to assess the ability of a single potential predictor variable to forecast future returns (e.g., Cochrane, 2011; Maio and Santa-Clara, 2012, hereafter MSC):

$$r_{t,t+q} = a_q + b_q x_t + \varepsilon_{t,t+q}, \tag{4.1}$$

where $r_{t,t+q} \equiv r_{t+1} + \cdots + r_{t+q}$ represents the continuously compounded return over q periods, that is, from $t+1$ to $t+q$; x_t is the value of the variable at time t whose predictive ability I want to assess; and $\varepsilon_{t,t+q}$ is a disturbance term (the forecasting error) with zero conditional mean, $E_t(\varepsilon_{t,t+q}) = 0$. The conditional expected return at time t can then be expressed as $E_t(r_{t,t+q}) = a_q + b_q x_t$. The forecasting power of x is assessed by regarding the degree of statistical significance of the slope coefficient, b_q, as well as by measuring the adjusted R^2 value of the regression. If the returns are unpredictable beyond a constant, that is, independent and identically distributed (i.i.d.), b_q is statistically indistinguishable from zero. Following MSC, I choose forecasting horizons of one month, three months, 12 months, 24 months, 36 months, 48 months, and 60 months ahead. The regressions are performed over the original sample period, January 1972 to June 2010, as well as over the two subsample periods January 1972 to December 1999 and January 2000 to June 2010, where q observations are lost in each respective q-horizon regression. By splitting the original sample into two subsamples, I seek to identify any structural changes over time, while the breakpoint is chosen to highlight the commodity boom of the 2000s. The regressions are conducted for each predictor proposed in Sec. 3.1.2 whose data series covers the respective period. Following MSC and others, I compute both Newey–West (1987) and Hansen–Hodrick (1980) t-ratios with q lags to assess the statistical significance of the regression coefficients. The q lags are selected to correct for the serial correlation in the regression residuals that are induced by the overlapping observations.

Next, I extend this single predictive regression model to a multiple-variable predictive regression model. Since my goal is not to test any existing theory that predetermines the right-hand side, it is unclear which of the variety of predictors I proposed should at once enter a multiple predictive regression. Instead, I seek to assess the marginal forecasting power of *each* candidate variable, conditional on *each* possible combination of all other variables. For this purpose, I employ the model selection procedure used by Bossaerts and Hillion (1999)

[29] This chapter is based on Sec. 3 of Lutzenberger (2014a).

and Zakamulin (2013), among others, for each horizon q. This procedure seeks to select the "best" regression model out of 2^N competing specifications of the following form:

$$r_{t,t+q} = \begin{cases} a_q + b_q' x_t + \varepsilon_{t,t+q}, & if\ n > 0, \\ a_q + \varepsilon_{t,t+q}, & if\ n = 0, \end{cases} \tag{4.2}$$

where $0 \leq n \leq N$ and x_t is a model-unique n-by-1 subvector of the vector of values at time t of all N candidate predictors. With regard to the large number of candidate predictors, I limit the set of competing regressions to those that include no more than seven independent variables to reduce the risk of overspecification and to keep computation times acceptable. Hence, I estimate each possible regression specification that includes no more than seven predictor variables, including the model with no predictors other than the constant, $n = 0$. The best model is then chosen according to a predefined model selection criterion. I choose the adjusted R^2 value for this purpose. If the returns are unpredictable beyond a constant, that is, i.i.d., the procedure should select the specification $n = 0$. I perform the model selection procedure for the same horizons and sample periods as for the single predictive regressions. However, so that the results across the three sample periods are comparable, I consider only those predictors whose data series cover the original sample period January 1972 to June 2010. Accordingly, I regard $N = 23$ potential predictive variables.

4.1.2. Out-of-sample prediction

The informative value of IS predictive regressions is not without controversy. Rather, some argue that the results might be spurious and might not hold OOS (e.g., Bossaerts and Hillion, 1999; Welch and Goyal, 2008; Zakamulin, 2013). For instance, Rapach et al. (2010) show that single predictive regressions with the aggregate earnings–price ratio, the aggregate dividend–payout ratio, the stock variance, the aggregate book-to-market ratio, the short-term risk-free rate, the long-term yield, the long-term return, the term structure spread, the default yield spread, the default return spread, and inflation have no OOS predictive power over the US equity premium, measured by quarterly S&P 500 excess returns over the period 1965 to 2005. For this reason, I also assess the OOS predictability of aggregate returns on commodity futures and employ an OOS predictive regression model that is based mainly on the OOS approaches used by Welch and Goyal (2008), Rapach et al. (2010), and Zakamulin (2013).

I first conduct individual OOS forecasts from single predictive regressions, as described by Rapach et al. (2010). Therefore, I start with a single predictive regression model for each candidate predictor, as formulated in Eq. (4.1), but I refrain from incorporating multiple horizons:

$$r_{t+1} = a + bx_t + \varepsilon_{t+1}, \tag{4.3}$$

where r_{t+1} is the return and ε_{t+1} is the disturbance term at time $t + 1$. OOS forecasts are then generated with a recursive (expanding) estimation window. I then split the total sample of observations for r_t and x_t into an initial IS period that consists of the first m observations and an OOS period that includes the last s observations. The initial OOS forecast of the return at time $m + 1$, based on a single predictor, is then computed as

$$\hat{r}_{m+1} = \hat{a}_m + \hat{b}_m x_m, \tag{4.4}$$

where $\hat{a}_m$ and $\hat{b}_m$ are the estimates of a and b, respectively, from Eq. (4.3), which are computed using observations of r_t from $t = 2$ to $t = m$ and of x_t from $t = 1$ to $t = m - 1$. In the next step, the estimation window is expanded by one period such that one obtains an estimate of the return at $m + 2$, $\hat{r}_{m+2}$, via observations of r_t from $t = 2$ to $t = m + 1$ and of x_t from $t = 1$ to $t = m$. I continue this procedure through the end of the OOS period and obtain a series of s OOS return forecasts based on a single predictor. This approach is conducted for each candidate predictor variable proposed in Sec. 3.1.2 and whose data series covers the original sample period January 1972 to June 2010.

I consider three different OOS periods: January 1980 to June 2010, January 1980 to December 1999, and January 2000 to June 2010. The IS period starts eight years prior to the beginning of the respective OOS period.[30]

In addition to these individual forecasts, I employ three combination forecasts proposed by Rapach et al. (2010): the mean, median, and trimmed mean forecasts. The mean combination method computes the arithmetic average of all individual forecasts of r_{t+1} made at time t to obtain another forecast of r_{t+1}. Analogously, the median combination forecast of r_{t+1} is the median of all individual forecasts of r_{t+1} made at t. Finally, the trimmed mean combination forecast computes the arithmetic average of all but the smallest and largest individual forecasts of r_{t+1} obtained at t.[31]

Finally, I use the OOS recursive forecasting procedure proposed by Zakamulin (2013). This procedure follows the individual forecasting method in Eqs. (4.3) and (4.4) but, instead of single predictive regressions, it conducts the model selection procedure described in Sec. 4.1.1. Hence, the first m observations are used to find the optimal (multiple-variable) predictive model to make the forecast $\hat{r}_{m+1}$. Following that, the IS period is expanded by one month and I repeat the procedure to find the best model, using data from $t = 1$ to $t = m + 1$, and compute the forecast $\hat{r}_{m+2}$. This procedure is continued through the end of the OOS period. As a result, I obtain a series of s OOS return forecasts, where each forecast is based on the best (multiple-variable) model, using only data prior to the month for which the forecast is made. To keep computation times manageable, I limit the set of potential predictors and, similar to Zakamulin (2013), consider only nine variables. Although this induces some degree of look-ahead bias, I use predictors that have performed relatively well IS: *D/E*, *SENT*, *LTY*, *DFR*, *CAY*, *M2*, *CCM_spot*, *C12CM_spot*, and *CVAR* (see Table 4 for their definitions).

I employ two measures to evaluate the individual, combination, and model selection OOS forecasts: the OOS R^2 statistic, R^2_{OS}, proposed by Campbell and Thompson (2008) and used by Rapach et al. (2010), among others, and the Henriksson–Merton (1981) test statistic, based on

30 I also considered IS periods that start at the beginning of the original sample, that is, in January 1972, following Rapach et al. (2010). However, the results were less convincing.

31 I also implemented more complex combining methods proposed by Rapach et al. (2010) that required a holdout OOS period to estimate the combining weights. However, the performance of these forecasts was not better than the simple schemes described above, thus confirming the results of Rapach et al.

the Henriksson–Merton test of directional accuracy (Henriksson and Merton, 1981; Pesaran and Timmermann, 1992) and employed by Zakamulin (2013).

The R_{OS}^2 is based on a series of moving historical averages of returns, $\bar{r}_{t+1} = \frac{1}{t}\sum_{j=1}^{t} r_j$, which are used as the benchmark for the respective return-forecasting method under evaluation. In particular, the statistic is computed as

$$R_{OS}^2 = 1 - \sum_{k=1}^{S}(r_{m+k} - \hat{r}_{m+k})^2 / \sum_{k=1}^{S}(r_{m+k} - \bar{r}_{m+k})^2. \tag{4.5}$$

The forecasting model under investigation, which generates the forecasts $\hat{r}_{t+1}$, outperforms the historical average forecast in terms of mean squared prediction errors if $R_{OS}^2 > 0$.

In a second step, I test whether the R_{OS}^2 is significantly greater than zero in two ways. First, I follow Rapach et al. (2010) and compute the *MSPE-adjusted* statistic proposed by Clark and West (2007). For this purpose, I first calculate

$$f_{t+1} = (r_{t+1} - \bar{r}_{t+1})^2 - [(r_{t+1} - \hat{r}_{t+1})^2 - (\bar{r}_{t+1} - \hat{r}_{t+1})^2] \tag{4.6}$$

and then regress the series of f_{t+1} on a constant and compute its t-statistic. The p-value for rejecting the null hypothesis $R_{OS}^2 \leq 0$ is then obtained using the standard normal distribution.

In addition to calculating the *MSPE-adjusted* statistic, I employ the non-parametric bootstrap method proposed by Zakamulin (2013). The null hypothesis, $R_{OS}^2 \leq 0$, corresponds to the null that returns are unpredictable and therefore i.i.d. The bootstrap method is used to estimate the sampling distribution of R_{OS}^2 under the conditions given by this null hypothesis. To be specific, after having computed the R_{OS}^2 for each forecasting model under investigation using the original time series of returns and predictors, I bootstrap the original time series to obtain random resamples of the returns and predictive variables. Accordingly, I resample the entire vector of returns and predictors at each time t to maintain the historical intratemporal correlations between these variables. I then compute R_{OS}^2 for each forecasting model, using the resampled time series of returns and predictive variables. This procedure is repeated numerous times and I count how many times R_{OS}^2 is above the R_{OS}^2 value obtained using the original time series to obtain empirical p-values for the null hypothesis, $R_{OS}^2 \leq 0$, against the alternative $R_{OS}^2 > 0$.

Finally, the Henriksson–Merton test statistic is computed as by Zakamulin (2013):

$$HM - S = Prob(\hat{r}_{t+1} > 0 | r_{t+1} > 0) + Prob(\hat{r}_{t+1} \leq 0 | r_{t+1} \leq 0), \tag{4.7}$$

where $Prob(\hat{r}_{t+1} > 0 | r_{t+1} > 0)$ is the conditional probability of obtaining a correct forecast, $\hat{r}_{t+1}$, of a positive return at time $t+1$, using the respective forecasting model under evaluation, given that the realized return at $t+1$, r_{t+1}, is positive. A forecasting model that is able to forecast the sign of the return generates a Henriksson–Merton test statistic that is greater than one, $HM - S > 1$. If the null that the return is unpredictable is true, the statistic should be unity, $HM - S = 1$. Following Zakamulin (2013), I test the null $HM - S \leq 1$

against the alternative hypothesis, $HM - S > 1$, by obtaining empirical p-values through the same non-parametric bootstrap method as that described above.

4.2. Methodology for Expected Returns in the Cross Section of European Stocks

In order to answer Research Questions 2 and 3, I estimate the CAPM as well as the three-factor model of Fama and French (1993), FF3; Carhart's (1997) four-factor model, C; the four-factor model of Pástor and Stambaugh (2003), PS; the five-factor model of Fama and French (1993), FF5; the unrestricted version of the ICAPM of Campbell and Vuolteenaho (2004), CV; the ICAPM of Hahn and Lee (2006), HL; the ICAPM of Petkova (2006), P; and the three-factor model of Koijen et al. (2010), KLVN, over the test assets SBM25 and SM25, which are constructed in Sec. 3.2.[32] On this basis, I evaluate whether these multifactor models meet the ICAPM criteria presented in Sec. 2.2.1.7. The following sections provide an overview of the empirical methodologies.

4.2.1. Predictive regressions

To evaluate ICAPM criterion 2, I have to test whether each model's state variables indeed forecast future investment opportunities. Second, to test ICAPM criterion 3a, I have to analyze whether the state variables are positively or negatively correlated with future market returns. Third, to test ICAPM criterion 3b, I have to assess the state variables' correlation with future market volatility. Therefore, I follow MSC and conduct various multivariate long-horizon predictive regressions for each model under investigation.

To test ICAPM criteria 2 and 3a, MSC conduct regressions with a future market return as the left-hand side variable and the current values of the *K* state variables that are included in the respective multifactor model as right-hand side variables. The regression results provide a picture of the joint forecasting power of the *K* state variables. Thus, the multivariate regressions help assess the marginal predictive role of each state variable for changes in future expected returns in the presence of the *K* - 1 remaining state variables. Specifically, for each multifactor model, I regress the continuously compounded market return over q periods (from $t+1$ to $t+q$), $r_{t,t+q}$, on the values of the corresponding state variables in t:

$$r_{t,t+q} = a_q + b_q' z_t + u_{t,t+q}, \tag{4.8}$$

where $r_{t,t+q} \equiv r_{t+1} + \cdots + r_{t+q}$, z_t is a *K*x1 vector of state variables and b_q is a *K*x1 vector of slope coefficients. In addition, $u_{t,t+q}$ is a forecasting error, which has a zero conditional mean. On this basis, the conditional expected market return at time t can be formulated as $E_t(r_{t,t+q}) = a_q + b_q' z_t$. The regression results indicate whether a specific state variable covaries positively or negatively with the expected future market return, conditional on the remaining state variables included in the multifactor model.

To assess ICAPM criteria 2 and 3b, MSC conduct regressions with the future realized stock market variance, $SVAR$, as the explained variable and the values of the *K* state variables in t as predictor variables:

[32] This chapter is based on Secs. 3 and 6 of Lutzenberger (2014b).

$$SVAR_{t,t+q} = a_q + b_q' z_t + v_{t,t+q}, \tag{4.9}$$

where $SVAR_{t,t+q}$ is the cumulative sum of $SVAR$ over q periods, $SVAR_{t,t+q} = SVAR_{t+1} + \cdots + SVAR_{t+q}$, and $v_{t,t+q}$ denotes a forecasting error, which has a conditional mean of zero. It follows that the conditional expectation of the future variance of market returns is $E_t(SVAR_{t,t+q}) = a_q + b_q' z_t$. The regression results indicate whether the state variables forecast positive or negative changes in future market volatility, given the values of the remaining state variables included in the multifactor model.[33] Following MSC, I use forecasting horizons of $q = 1, 3, 12, 24, 36, 48,$ and 60 months for all regressions, and compute Newey–West (1987) asymptotic standard errors with q lags to correct for the serial correlation in the residuals resulting from the overlapping returns.

4.2.2. Generalized method of moments

Using a one-stage generalized method of moments (GMM) procedure following Hansen (1982), MSC estimate each model in expected return–covariance form. The advantage of the GMM approach and the estimation of the models in expected return–covariance form is that one can directly assess the market (covariance) risk price and see whether its value is economically plausible as an estimate of the relative risk aversion of the representative investor (ICAPM criterion 1). The first-stage GMM procedure is conceptually equivalent to a cross-sectional ordinary least squares (OLS) regression of the assets' average excess returns on the covariances between the returns and factors. However, the resulting GMM standard errors correct for the fact that the means of the factors are estimated. I refer to Appendix A for the GMM formulas. Two measures are used to assess the goodness of fit of the multifactor models in the cross section of expected excess returns. First, I compute the mean absolute pricing error:

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |\hat{\alpha}_i|, \tag{4.10}$$

where $\hat{\alpha}_i, i = 1, \ldots, N$, denotes the pricing errors of the N test assets. The second measure is the cross-sectional OLS coefficient of determination:

$$R_{OLS}^2 = 1 - \frac{Var_N(\hat{\alpha}_i)}{Var_N(\bar{R}_i)}, \tag{4.11}$$

where $\bar{R}_i$ is the average excess return of test asset i and $Var_N(\cdot)$ denotes the cross-sectional variance.

Implementing the multifactor models requires estimation of the innovations in the state variables, which are generally not just simple changes. The original papers suggest different time-series processes to proxy for the innovations. Hahn and Lee (2006) use first differences in the state variables, while Campbell and Vuolteenaho (2004), Petkova (2006), and Koijen et

[33] As a third aspect, MSC also consider simultaneous variations in the first two moments of aggregate returns by estimating the conditional Sharpe ratio for each model and at each horizon. However, to save space, I limit the study to a separate analysis of the first two moments.

al. (2010) specify a first-order vector autoregressive VAR(1) process that their state variables follow. I follow MSC and use first differences as proxies for the state variable innovations,

$$\Delta z_{t+1} \equiv z_{t+1} - z_t, \tag{4.12}$$

and the main discussion is based on the empirical results obtained with this proxy. Different proxies are considered in Sec. 4.2.3.

4.2.3. Sensitivity analysis

I conduct some of the robustness checks performed by MSC to assess how sensitive my empirical results for the European stock market are to changes in the underlying empirical methodology. First, I employ alternative proxies for the state variables associated with the size factor SMB and the value factor HML in the predictive regressions. These alternative proxies are constructed in the same way as the liquidity factor CL and the momentum premium $CUMD$. Second, I replace the stock market variance $SVAR$ by three alternative volatility measures in the volatility-predicting regressions, that is, the volatility measure proposed by Beeler and Campbell (2012), the proxy suggested by Bansal et al. (2005), and the monthly squared continuously compounded market return. Third, I include an intercept in each model's pricing equation. Fourth, I estimate all models with second-stage GMM. Fifth, I add the excess returns on seven artificial German government bonds with residual maturities of one year, two years, five years, seven years, 10 years, 15 years, and 20 years to both sets of testing assets, for a total of 33 testing assets (i.e., 25 stock portfolios, the equity premium, and seven bonds) in each test. The bond returns are obtained using the term structure of interest rates published by Deutsche Bundesbank. Sixth, I exclude the market excess return from the set of testing assets. Seventh, I estimate all multifactor models in expected return–beta form with the time-series/cross-sectional regression approach employed by, for example, Brennan et al. (2004) and Cochrane (2005, Sec. 12.2). I thereby conduct both OLS as well as generalized least squares cross-sectional regressions. Eighth, I estimate each model by firstly orthogonalizing each risk factor with respect to the market excess return. Finally, I use different proxies for the state variable innovations within the ICAPM specifications, that is, I specify first-order autoregressive AR(1) processes as well as a VAR(1) process for each model to derive the state variable innovations. A more detailed presentation of these methodologies is given in Appendix B.

4.3. Methodology for the Cost of Equity Capital for European Industries

To investigate Research Questions 4 and 5, I formulate the CAPM and the multifactor models FF3, C, PS, FF5, CV, HL, P, and KLVN in expected return–beta form:

$$E(R_i^e) = \beta_i' \lambda, \tag{4.13}$$

where $E(R_i^e)$ is the expected excess return on asset i over the risk-free rate, λ is a K-by-1 vector of risk premiums that are associated with the model's K risk factors, and β_i is a K-by-1 vector of asset i's risk loadings on the K risk factors.[34] The risk loadings are the slope coefficients of the time-series regression

$$R_{i,t}^{e} = a_i + \beta_i' f_t + \varepsilon_{i,t}, \tag{4.14}$$

where $R_{i,t}^{e}$ is the excess return on asset i at time t and f_t is a K-by-1 vector of the K risk factors. These risk factors are summarized in Table 2. Moreover, a_i is a constant and $\varepsilon_{i,t}$ is the regression's residual at time t. If the factors are excess returns themselves, that is, traded portfolios, then the risk premiums are equal to the expected values of the risk factors, $\lambda = E(f)$. In this case, a_i represents the pricing error of asset i, that is, that part of the asset's return that remains unexplained by the pricing model. The model is correctly specified if the pricing errors of all assets are jointly zero (e.g., Cochrane, 2005, Sec. 12; Goyal, 2012).

4.3.1. Estimation of risk loadings

Following Fama and French (1997, hereafter FF97) and Gregory and Michou (2009, hereafter GM09), I first estimate, for each industry and each model under investigation, regression (4.14) over the full sample period, August 1990 to December 2011, via OLS, implicitly assuming that the true factor risk loadings are constant through time. However, FF97 note that the slope standard errors obtained from full-period regressions might be misleading because of the assumption of constant true factor risk loadings. To assess whether the true risk loadings vary over time, I follow FF97 and GM09 and, second, conduct 60-month rolling regressions to estimate the implied standard deviations of the true factor risk loadings. In doing so, I assume that the true loadings follow stationary processes. Following FF97, the implied standard deviation $\sigma(True)$ of a given industry's true risk loading is estimated as

$$\hat{\sigma}(True) = [\hat{\sigma}^2(Time\ Series) - \hat{\sigma}^2(Estimation\ Error)]^{1/2}, \tag{4.15}$$

where $\hat{\sigma}^2(Time\ Series)$ is the time-series variance of the industry's rolling regression slope estimates and $\hat{\sigma}^2(Estimation\ Error)$ is the mean of the estimation error variances (squared standard errors) of the industry's rolling regression slope estimates. The estimation error variances are heteroskedasticity consistent, following White (1980). Following FF97, I set an industry's implied standard deviation equal to zero if the average estimation error variance exceeds the time-series variance of the industry's rolling regression slope estimates. Finally, I compute an upper (lower) bound of the typical industry's true loadings on a given risk factor by adding (subtracting) twice the slope coefficient's average implied standard deviation across all industries to (from) the average slope coefficient across all months and industries. The rolling regressions use data over the whole sample period August 1990 to December 2011, but 59 months are lost when I compute the implied standard deviations.

Another way to track the time variation of risk loadings is by formulating the investigated asset pricing models in conditional form (e.g., Jagannathan and Wang, 1996; Cochrane, 2005, Ch. 8), as done by FF97 and GM09 for the CAPM and FF3. However, a wide variety of candidate state variables to track the wandering risks exists, resulting in many different

[34] This chapter is based on Sec. 2 and Sec. 4 of Lutzenberger (2014c).

possible alternatives for formulating these models. Therefore, to save space, I limit myself to rolling regressions.

A financial manager might want to account for the time variation of true factor risk loadings in some way when estimating an industry's cost of equity capital (CE). One could argue that, in doing so, the manager should take into account the maturities of a project's cash flows. FF97 note that rolling or conditional regressions are a priori likely to be the appropriate methods for estimating the cost of equity of near-term cash flows because these cash flows need to be discounted with the current CE, which requires the current true risk loadings. Whether more distant cash flows should be discounted using current or long-term average risk loadings depends on the behavior of the true risk loadings over time. If the true risk loadings are mean reverting, the manager may prefer the average risk loadings estimated by static regressions. In contrast, the current true risk loadings estimated by rolling or conditional regressions are probably better candidates for more distant cash flows when the true risk loadings are generated by a random walk.

To test these suggestions, I follow FF97 and conduct various forecasts of short- and long-term CEs using full-period and rolling estimates. I include an intercept in the time-series regressions to estimate the risk loadings, but I drop the estimated intercept when I conduct the return forecasts. Moreover, I limit myself to the CAPM, FF3, and C because of the additional difficulties implied by the time-series intercepts of the remaining models not being their pricing errors. Furthermore, for simplicity, I do not conduct the Bayes shrinkage method suggested by FF97. To be specific, I compute monthly CE forecast errors from $t = 60 + q$ to $t = T$, where $t = 60$ corresponds to August 1995, $t = T$ is December 2011, and $q = 1, 12, 24, 36, 48, 60$ months is the forecasting horizon. In the first step, I conduct rolling regressions for each of the three models and for each industry. The forecasting error at t from these regressions is computed as the industry's realized return at t minus the vector of realized factors at t times the vector of factor slopes that were estimated from $t - q - 59$ to $t - q$. Thus, the intercept from the rolling regressions is not considered. I then compute, for each model and each forecasting horizon, the average mean absolute forecast error, as well as the average (across all industries) standard deviation of the forecast errors. In a second step, I estimate static (full-period) regressions from $t = 60 + q$ to $t = T$. The corresponding forecast error at t is then computed as the regression's residual at t plus its intercept. For the static regressions, I adjust both the average mean absolute forecast errors and the average standard deviations of the forecast errors for degrees of freedom by multiplying by $\frac{T-59-q-1}{T-59-q-K-1}$, where K is the number of factors included in the model in question.

Finally, I follow GM09 and additionally examine the errors from estimating the observed risk loadings 60 months ahead, using rolling regressions that include 60 months of past returns. This test also allows me to examine the models with non-traded factors. In particular, I start with computing a factor's slope prediction error in month t as the slope of a regression involving excess returns and factors from $t + 1$ to $t + 60$ minus the slope of a regression with excess returns and factors from $t - 59$ to t. Then I calculate the mean and the average standard deviation of these prediction errors, as well as the mean of the absolute prediction errors. I repeat this procedure for each factor and each of the nine models under investigation, that is, for the CAPM and all eight multifactor models, and I use the original sample period August 1990 to December 2011, losing 119 months when I compute the prediction errors.

4.3.2. Estimation of risk premiums

FF97 note that the estimation of factor risk loadings is only a small part of the cost of equity assessment. They claim that uncertainty about factor risk premiums is even more important. Therefore, I devote a brief section to factor risk premiums, especially differences between the CAPM, FF3, and C, which include only traded factors, and PS, FF5, HL, P, CV, and KLVN, which include both traded and non-traded factors.

If the factor is an excess return, one can use the sample mean $E_T(\cdot)$ of the factor as an estimate of its risk premium (e.g., Cochrane, 2005, Sec. 12.1):

$$\hat{\lambda} = E_T(f), \tag{4.16}$$

which applies to the excess market return RM, the size factor SMB, the value factor HML, and the momentum factor UMD.

Estimation of the risk premiums for non-traded factors requires "something like a cross-sectional regression" (Cochrane, 2005, p. 244). I therefore choose the method suggested by Fama and MacBeth (1973).

Lewellen et al. (2010) note that one test of an asset pricing model with traded factors is whether the risk premiums estimated on the basis of time-series and cross-sectional regressions are statistically indistinguishable (see also Goyal, 2012). However, some of the models I investigate might fail this test and the risk premiums associated with the traded factors estimated by cross-sectional regressions might therefore deviate from the factors' sample means, at least to some degree. However, I initially want to keep the historical averages as risk premiums for the traded factors when I later estimate the industries' CEs using the models that include non-traded factors. Only in a second step do I want to estimate the all-factor risk premiums using cross-sectional regressions.

I therefore first impose the null hypothesis that the risk premiums for RM, SMB, HML, and UMD are equal to their means and move these factors to the left-hand side of the Fama–MacBeth cross-sectional regressions. This approach is sometimes used to reduce the errors in variables problem (e.g., Brennan et al., 1998; Goyal, 2012). In particular, I conduct the following monthly cross-sectional regressions for each model that contains non-traded factors:

$$R_{i,t}^{e} - (\hat{\beta}_{i,traded}')f_{t,traded} = (\hat{\beta}_{i,non-traded}')\lambda_{t,non-traded} + \varepsilon_t, \tag{4.17}$$

where $\hat{\beta}_{i,traded}$ ($\hat{\beta}_{i,non-traded}$) is an M-by-1 (N-by-1) vector of the estimated risk loadings on the M (N) traded (non-traded) factors included in the model. The risk loadings are estimated using either full-period multiple time-series regressions or 60-month rolling multiple time-series regressions (as in Sec. 4.3.1). Moreover, $f_{t,traded}$ is an M-by-1 vector of the model's M traded factors in month t. I obtain the risk premiums associated with the non-traded factors as the means of the monthly estimates of $\lambda_{t,non-traded}$, that is, of the estimated slope coefficients of regression (4.17):

$$\hat{\lambda}_{non-traded} = E_T(\hat{\lambda}_{t,non-traded}). \tag{4.18}$$

These regressions are conducted over the full sample period, August 1990 to December 2011, and 59 months are lost when the factor loadings are estimated by rolling regressions.

In the next step, I remove the null hypothesis that the risk premiums for traded factors are equal to the factor means. I then conduct Fama–MacBeth regressions with the estimated slopes on both traded and non-traded factors as independent variables to estimate all factor risk premiums, again considering the full sample period, August 1990 to December 2011, where 59 months are lost with rolling regressions. To be specific, I run the following monthly cross-sectional regressions for all nine models under investigation:

$$R_{i,t}^{e} = \hat{\beta}_i' \lambda_t + \varepsilon_t, \tag{4.19}$$

where $\hat{\beta}_i$ is a *K*-by-1 vector of the estimated loadings on the *K* risk factors included in the model. These loadings are again estimated using either full-period multiple time-series regressions or 60-month rolling multiple time-series regressions. The risk premiums associated with the *K* risk factors are then obtained as the means of the monthly estimates of λ_t (which are the estimated slope coefficients of regression (4.19)):

$$\hat{\lambda} = E_T(\hat{\lambda}_t). \tag{4.20}$$

4.3.3. Estimation of the industry costs of equity capital

Finally, I compute the CEs (more precisely, the industry-specific risk-premiums, i.e., their expected returns in excess of the risk-free rate) for the different approaches to estimate the risk loadings and risk premiums discussed above. Specifically, I estimate the CEs for all industries in four different ways, by combining two methods for the estimation of risk loadings with two methods for estimating risk premiums: risk loadings estimated from either (a) full-period time-series regressions or (b) 60-month rolling regressions combined with either (a) risk premiums estimated using factor sample means (for traded factors) and the Fama–MacBeth approach as in Eqs. (4.17) and (4.18) (for non-traded factors) or (b) risk premiums estimated with Fama–MacBeth regressions for all factors, as in Eqs. (4.19) and (4.20).

5. Empirical Results

5.1. Results for Expected Returns on Commodity Futures in the Time Series

The following sections present my empirical results with regard to the question whether aggregate returns on commodity futures are predictable in the time series (Research Question 1).[35]

5.1.1. In-sample results

Table 10 shows the results for the in-sample (IS) single predictive regressions. With three sample periods, 23 or more candidate predictors for each sample period, and seven forecasting horizons, there is reason to limit the results shown to variables that significantly predict future returns at least one horizon at the 5% level (as indicated by either Newey–West (1987) or Hansen–Hodrick (1980) t-ratios). To save even more space, the Hansen–Hodrick t-ratios are shown only for those variables that would have not been included in the table if regarding only the Newey–West values.

Note that *SVAR*, *B/M*, *LTR*, *DEF*, *INFL*, *I/K*, *IP*, *GDP*, *CCM_spot*, and *C12CM_spot* are unable to forecast either short- or long-term returns within the original sample period. Hence, the results for these variables are only available upon request. Nonetheless, the regression coefficients of the other variables I consider are significantly different from zero. On the one hand, *DY*, *E/P*, *CRMRF*, *CL*, *SENT*, *RF*, *LTY*, and *CAY* consistently forecast negative returns; on the other hand, *D/E*, *TERM*, *DFR*, and *M2* predict positive returns. Additionally, the predictive sign of *CVAR* is negative for horizons of one month and three months (but without significance) and becomes positive for longer horizons. Consequently, aggregate returns on commodity futures seem to be predictable IS beyond a constant and are therefore not independent and identically distributed (i.i.d.). Observe that the predictive power of most variables increases with the horizon according to the adjusted R^2 values. This is due to the high persistence of most predictors, that is, their slow movement, as indicated by their autocorrelation coefficients above 0.9 (Cochrane, 2005, Sec. 20.1).

For horizons of one month to 12 months, *SENT* shows the highest forecasting power according to the t-statistics and R^2. Thus, out of the set of candidate predictors considered, a high sentiment level of stock market investors seems to be the "best" single predictor for low subsequent short-horizon aggregate returns on commodity futures. Regarding horizons of 24 months and longer, we see that a high level of *RF,* followed by *LTY* and *TERM,* does a good job of indicating low future long-horizon returns. Consequently, much forecastability of long-horizon aggregate returns on commodity futures appears to be due to the current level and slope of the yield curve.

[35] This chapter is based on Sec. 4 of Lutzenberger (2014a).

Table 10. Single predictive regressions.

This table displays the results for single predictive regressions for the monthly continuously compounded return on an equal-weighted portfolio of 27 commodity futures at horizons q = 1, 3, 12, 24, 36, 48, and 60 months ahead. The returns do not include the returns on collateral from transacting in futures contracts. The original sample period is January 1972 to June 2010 (Panel A), the two subsample periods are January 1972 to December 1999 (Panel B) and January 2000 to June 2010 (Panel C), respectively, and q observations are lost in each of the respective q-horizon regressions. The first line corresponding to each model reports the slope estimates. Line 2 reports the Newey–West t-ratios (in parentheses) and, where required, line 3 shows the Hansen–Hodrick t-ratios (in square brackets), both computed with q lags. Italic, underlined, and boldfaced t-statistics denote statistical significance according to the standard normal distribution at the 10%, 5%, and 1% levels, respectively. The last line, R^2 (%), shows the values of the adjusted coefficient of determination (%). The table reports only those regression models whose slope coefficients are significant for at least one horizon according to either the Newey–West or Hansen–Hodrick t-ratio.

Panel A: January 1972 to June 2010

Predictor	$q = 1$	$q = 3$	$q = 12$	$q = 24$	$q = 36$	$q = 48$	$q = 60$
DY	0.00	-0.01	-0.03	-0.07	-0.16	-0.25	-0.32
	(-0.55)	(-0.66)	(-0.46)	(-0.74)	(-1.23)	(-1.64)	(-2.09)
R^2 (%)	-0.16	-0.03	0.10	1.10	5.36	10.43	14.06
E/P	0.00	-0.01	-0.05	-0.13	-0.25	-0.34	-0.41
	(-0.56)	(-0.84)	(-1.05)	(-1.34)	(*-1.95*)	(-2.46)	(**-2.88**)
R^2 (%)	-0.13	0.17	1.10	3.86	12.00	19.01	23.18
D/E	0.00	0.01	0.06	0.25	0.35	0.43	0.51
	(0.23)	(0.57)	(1.53)	(1.20)	(1.51)	(*1.87*)	(*1.73*)
	[0.22]	[0.50]	[1.28]	[1.05]	[*1.70*]	[**2.65**]	[2.27]
R^2 (%)	-0.21	-0.09	0.75	2.97	5.01	5.64	6.06
CRMRF	-0.01	-0.03	-0.18	-0.34	-0.36	-0.27	-0.24
	(-1.53)	(*-1.70*)	(-2.49)	(**-3.27**)	(**-3.39**)	(-2.34)	(-1.59)
R^2 (%)	0.40	1.29	8.14	13.81	13.00	6.08	3.83
CL	-0.01	-0.02	-0.07	-0.12	-0.15	-0.14	-0.11
	(*-1.70*)	(*-1.83*)	(-2.06)	(-2.33)	(**-2.66**)	(-1.99)	(-1.23)
R^2 (%)	0.57	1.56	4.39	6.13	8.00	5.83	3.02
SENT	-0.01	-0.02	-0.07	-0.11	-0.10	-0.08	-0.05
	(**-2.93**)	(**-3.09**)	(**-3.05**)	(**-2.86**)	(-2.23)	(-1.47)	(-0.79)
R^2 (%)	1.96	4.59	10.89	11.57	8.25	4.63	1.20
RF	-0.12	-0.39	-1.84	-4.08	-6.17	-6.93	-6.59
	(-1.54)	(*-1.75*)	(-2.31)	(**-4.07**)	(**-6.94**)	(**-5.86**)	(**-3.96**)
R^2 (%)	0.58	1.87	7.21	16.55	32.58	33.96	25.61
LTY	-0.20	-0.62	-2.49	-4.34	-5.97	-7.07	-7.72
	(-2.24)	(-2.40)	(-2.27)	(-2.47)	(**-2.82**)	(**-2.59**)	(-2.52)
R^2 (%)	1.04	3.02	8.45	12.12	19.48	21.58	20.65
TERM	0.01	0.03	1.02	4.78	8.71	10.06	7.98
	(0.03)	(0.08)	(0.60)	(2.15)	(**3.39**)	(**3.34**)	(*1.82*)
R^2 (%)	-0.22	-0.21	0.35	5.82	17.37	18.34	9.49
DFR	0.38	0.59	1.37	1.79	1.17	1.72	1.62
	(1.39)	(1.17)	(2.25)	(*1.80*)	(1.13)	(*1.67*)	(1.59)
R^2 (%)	1.30	0.73	0.61	0.23	-0.08	0.03	-0.04
CAY	-0.23	-0.65	-2.03	-2.45	-2.79	-3.57	-4.48
	(-2.44)	(-2.30)	(-1.40)	(-0.97)	(-1.02)	(-1.06)	(-1.02)
R^2 (%)	1.20	2.81	4.79	3.42	3.72	4.60	5.80
M2	0.05	0.16	0.58	0.87	0.98	1.04	1.09
	(2.04)	(1.97)	(1.52)	(1.35)	(1.37)	(1.28)	(1.13)
R^2 (%)	0.69	1.85	4.36	4.96	5.56	5.14	4.71
CVAR	-4.43	-6.33	21.41	9.94	20.30	37.93	76.01
	(-1.04)	(-0.67)	(0.96)	(0.31)	(0.70)	(1.26)	(*1.72*)
	[-0.94]	[-0.59]	[1.11]	[0.41]	[0.81]	[1.96]	[2.06]
R^2 (%)	0.18	0.00	0.19	-0.20	-0.13	0.07	0.80

Table 10 (continued)

Panel B: January 1972 to December 1999

Predictor	$q = 1$	$q = 3$	$q = 12$	$q = 24$	$q = 36$	$q = 48$	$q = 60$
DY	0.00	-0.01	-0.04	-0.15	-0.42	-0.68	-0.79
	(-0.36)	(-0.54)	(-0.31)	(-0.57)	(-1.29)	(-1.99)	(-2.03)
R^2 (%)	-0.26	-0.07	0.04	1.70	10.74	21.04	21.92
E/P	0.00	-0.01	-0.05	-0.17	-0.33	-0.46	-0.48
	(-0.27)	(-0.54)	(-0.56)	(-0.90)	(-1.59)	(-2.02)	(*-1.77*)
R^2 (%)	-0.28	-0.08	0.61	3.34	11.86	19.83	19.06
CRMRF	-0.01	-0.03	-0.17	-0.35	-0.36	-0.30	-0.37
	(-1.25)	(-1.14)	(*-1.70*)	(-2.09)	(-2.20)	(*-1.85*)	(*-1.75*)
R^2 (%)	0.48	1.07	5.79	9.51	8.64	4.86	6.07
CL	-0.01	-0.02	-0.06	-0.10	-0.10	-0.09	-0.07
	(-1.44)	(-1.49)	(*-1.74*)	(*-1.76*)	(-1.62)	(-1.11)	(-0.75)
	[-1.37]	[-1.34]	[-1.43]	[*-1.74*]	[-2.27]	[-1.38]	[**-2.86**]
R^2 (%)	0.55	1.45	3.63	4.04	3.54	1.96	0.84
SENT	-0.01	-0.02	-0.07	-0.12	-0.12	-0.12	-0.09
	(-2.09)	(-2.34)	(**-2.65**)	(**-2.98**)	(**-3.15**)	(-2.29)	(-1.55)
R^2 (%)	1.29	3.70	10.21	14.10	14.23	10.96	5.27
RF	-0.22	-0.65	-2.56	-4.92	-6.88	-7.92	-7.17
	(-2.36)	(**-2.74**)	(-2.42)	(**-2.73**)	(**-3.76**)	(**-3.86**)	(**-2.80**)
R^2 (%)	1.64	4.30	10.92	18.12	32.34	37.05	26.47
LTY	-0.35	-1.07	-4.14	-7.03	-8.56	-9.67	-9.29
	(**-3.26**)	(**-3.37**)	(-2.32)	(-2.29)	(**-2.67**)	(-2.51)	(-2.01)
R^2 (%)	2.79	7.46	17.31	20.96	27.47	29.62	22.97
LTR	-0.11	-0.09	-0.07	0.15	0.38	0.59	1.06
	(-1.48)	(-0.60)	(-0.21)	(0.28)	(0.60)	(0.82)	(1.62)
	[-1.43]	[-0.57]	[-0.18]	[0.31]	[0.78]	[1.00]	[2.20]
R^2 (%)	0.33	-0.18	-0.30	-0.30	-0.21	-0.08	0.37
TERM	0.01	-0.04	0.25	3.05	6.90	8.47	7.27
	(0.03)	(-0.06)	(0.12)	(1.07)	(2.15)	(2.46)	(1.41)
R^2 (%)	-0.30	-0.30	-0.28	1.88	10.03	13.40	8.56
DEF	-1.05	-3.17	-9.75	-17.27	-18.96	-16.74	-14.49
	(-2.14)	(-2.03)	(-1.40)	(*-1.69*)	(*-1.75*)	(-1.15)	(-0.96)
R^2 (%)	1.00	2.91	4.48	6.25	6.60	4.21	2.46
CP	-0.20	-0.53	-1.33	0.22	1.40	1.93	2.43
	(*-1.86*)	(*-1.92*)	(-1.40)	(0.14)	(0.68)	(0.93)	(0.97)
	[*-1.80*]	[*-1.71*]	[-1.27]	[0.15]	[0.92]	[2.12]	[1.30]
R^2 (%)	1.01	2.29	2.24	-0.29	0.83	1.59	2.33
INFL	0.08	0.12	-3.51	-16.18	-27.32	-32.41	-35.82
	(0.09)	(0.05)	(-0.45)	(-1.22)	(*-1.84*)	(-1.97)	(*-1.67*)
R^2 (%)	-0.30	-0.30	0.01	2.71	7.36	8.99	9.43
CAY	-0.28	-0.86	-3.27	-4.62	-3.88	-3.32	-3.26
	(-2.07)	(-2.10)	(-1.64)	(-1.34)	(-0.98)	(-0.68)	(-0.54)
R^2 (%)	1.83	5.05	11.82	10.52	6.69	4.01	3.16
M2	0.05	0.16	0.57	0.82	0.80	0.74	0.84
	(2.07)	(*1.95*)	(1.47)	(1.26)	(1.14)	(1.00)	(0.93)
R^2 (%)	0.93	2.49	5.37	5.11	4.19	2.73	2.53
C12CM_spot	0.06	0.18	0.54	0.47	0.24	0.12	0.21
	(2.35)	(**2.93**)	(2.28)	(1.54)	(0.59)	(0.26)	(0.39)
R^2 (%)	3.23	7.49	9.95	3.03	0.48	-0.17	0.11
CVAR	1.33	5.62	34.47	14.95	31.51	48.03	84.18
	(0.28)	(0.53)	(1.19)	(0.47)	(1.19)	(*1.89*)	(2.20)
R^2 (%)	-0.27	-0.15	0.60	-0.25	-0.02	0.29	1.36

Table 10 (continued)

Panel C: January 2000 to June 2010

Predictor	$q = 1$	$q = 3$	$q = 12$	$q = 24$	$q = 36$	$q = 48$	$q = 60$
E/P	-0.01	-0.02	-0.07	-0.32	-0.57	-0.73	-0.92
	(-0.85)	(-1.30)	(-1.17)	(*-1.70*)	(**-4.59**)	(**-5.97**)	(**-8.57**)
R^2 (%)	-0.39	0.28	1.02	12.45	41.90	58.03	68.50
D/E	0.00	0.01	0.06	0.44	0.77	0.78	0.88
	(0.32)	(0.63)	(*1.74*)	(2.28)	(**5.45**)	(**4.57**)	(**4.06**)
R^2 (%)	-0.73	-0.34	1.46	17.53	56.19	55.08	63.46
SVAR	-1.55	-2.33	0.99	3.75	28.70	37.82	46.40
	(-2.19)	(-1.96)	(0.47)	(0.23)	(*1.81*)	(2.10)	(2.54)
R^2 (%)	4.88	2.32	-0.77	-0.85	7.91	14.41	21.34
B/M	-0.01	0.11	0.87	0.76	-0.58	-1.45	-1.03
	(-0.22)	(0.77)	(2.48)	(1.42)	(-0.70)	(-1.43)	(-1.32)
R^2 (%)	-0.78	-0.13	7.51	2.92	1.35	14.24	5.02
CRMRF	-0.01	-0.04	-0.33	-0.52	-0.34	0.00	0.07
	(-1.05)	(*-1.67*)	(**-3.34**)	(**-4.70**)	(**-4.02**)	(0.01)	(0.49)
R^2 (%)	-0.31	1.89	30.98	50.01	20.48	-1.32	-0.53
CL	-0.02	-0.06	-0.33	-0.51	-0.62	-0.32	0.16
	(-1.20)	(-1.37)	(-2.28)	(**-3.72**)	(**-4.38**)	(-2.47)	(0.99)
R^2 (%)	0.28	3.41	23.41	40.28	49.60	11.28	0.89
SENT	-0.02	-0.04	-0.12	-0.10	-0.02	0.02	0.05
	(**-4.07**)	(**-4.81**)	(**-3.62**)	(*-1.87*)	(-0.44)	(0.36)	(1.11)
R^2 (%)	5.41	9.77	18.29	7.42	-0.59	-1.06	0.66
RF	-0.01	-0.18	-2.71	-8.36	-11.20	-7.32	-0.75
	(-0.03)	(-0.35)	(-1.34)	(**-4.72**)	(**-4.81**)	(**-2.71**)	(-0.35)
R^2 (%)	-0.81	-0.68	4.89	32.04	59.47	22.27	-1.32
LTY	-0.06	-0.41	-2.53	0.04	-2.02	13.81	10.51
	(-0.10)	(-0.23)	(-0.37)	(0.00)	(-0.25)	(1.18)	(1.02)
	[-0.09]	[-0.19]	[-0.39]	[0.00]	[-0.46]	[1.02]	[2.01]
R^2 (%)	-0.81	-0.74	-0.28	-1.00	-0.95	7.43	2.98
LTR	-0.02	-0.10	0.04	-0.25	0.67	0.57	-0.33
	(-0.18)	(-0.69)	(0.14)	(-0.57)	(0.83)	(0.97)	(-1.11)
	[-0.20]	[-0.64]	[0.30]	[-0.57]	[1.05]	[1.04]	[**-3.67**]
R^2 (%)	-0.79	-0.69	-0.89	-0.93	-0.61	-0.93	-1.45
TERM	0.00	0.19	3.16	9.72	12.74	11.34	2.64
	(0.00)	(0.36)	(1.40)	(**4.48**)	(**7.26**)	(**3.54**)	(1.28)
R^2 (%)	-0.81	-0.72	5.08	37.44	66.63	42.40	0.42
DEF	-0.38	-0.16	4.66	23.15	52.06	72.39	80.87
	(-0.37)	(-0.06)	(1.24)	(0.65)	(**3.39**)	(**4.81**)	(**4.03**)
R^2 (%)	-0.61	-0.82	0.76	3.10	16.26	34.78	46.04
DFR	0.52	0.79	1.43	4.25	1.79	1.65	1.97
	(1.27)	(0.97)	(1.62)	(**2.85**)	(*1.69*)	(1.41)	(**3.08**)
R^2 (%)	4.74	2.29	1.05	3.41	-0.61	-0.85	-0.89
INFL	-0.06	-1.59	-10.20	-6.67	-9.19	-12.12	-12.54
	(-0.06)	(-0.55)	(*-1.78*)	(-0.76)	(*-1.81*)	(**-3.75**)	(**-3.64**)
R^2 (%)	-0.81	-0.24	4.11	-0.04	0.62	1.38	0.42
I/K	-1.13	-4.21	-27.18	-35.40	-23.22	-8.84	2.00
	(-1.52)	(-2.09)	(**-4.40**)	(**-3.99**)	(**-2.63**)	(-0.97)	(0.17)
R^2 (%)	0.23	2.58	20.98	22.09	9.63	0.35	-1.47

Table 10 (continued)

Panel C (continued): January 2000 to June 2010							
Predictor	$q = 1$	$q = 3$	$q = 12$	$q = 24$	$q = 36$	$q = 48$	$q = 60$
CAY	-0.25	-0.31	3.21	9.96	12.87	15.70	19.04
	(-1.05)	(-0.55)	(1.11)	(**3.87**)	(**10.35**)	(**7.44**)	(**5.95**)
R^2 (%)	-0.17	-0.58	4.94	39.66	62.53	64.35	64.48
IP	-0.02	-0.13	-0.97	-1.63	-0.76	0.62	0.94
	(-0.54)	(-1.41)	(**-4.51**)	(**-5.81**)	(-2.04)	(0.74)	(1.01)
R^2 (%)	-0.65	0.72	12.77	19.16	3.65	2.04	5.92
M2	0.08	0.32	1.42	3.43	6.29	8.85	11.08
	(0.63)	(0.86)	(0.95)	(1.60)	(**6.96**)	(**18.16**)	(**9.91**)
R^2 (%)	-0.56	0.12	3.12	15.60	53.74	78.90	56.18
GDP	-0.01	-0.11	-1.32	-4.10	-7.50	-7.20	-3.65
	(-0.10)	(-0.54)	(-1.62)	(-2.02)	(**-5.23**)	(**-5.08**)	(**-2.87**)
R^2 (%)	-0.81	-0.55	4.19	16.88	61.20	54.84	9.23
USD	-0.01	-0.01	0.09	0.36	0.58	0.86	1.02
	(-0.46)	(-0.18)	(0.29)	(0.91)	(*1.70*)	(2.49)	(**3.13**)
R^2 (%)	-0.68	-0.76	-0.20	7.80	18.87	32.51	31.93
CVAL	0.05	0.13	0.49	0.57	0.07	-0.71	-0.66
	(*1.84*)	(*1.68*)	(**2.70**)	(**3.70**)	(0.31)	(-1.57)	(*-1.71*)
R^2 (%)	1.56	3.91	13.48	13.18	-0.91	23.17	16.44
CMOM	-0.03	-0.06	0.13	0.57	0.42	1.54	1.34
	(-1.10)	(-1.22)	(0.30)	(0.75)	(0.75)	(2.51)	(**3.06**)
R^2 (%)	-0.20	-0.05	-0.45	4.87	1.59	32.32	22.09
CCM_spot	-0.01	-0.03	-0.15	-0.35	-0.45	-0.64	-0.93
	(-0.44)	(-0.62)	(-0.75)	(-1.53)	(-2.49)	(**-3.70**)	(**-6.19**)
R^2 (%)	-0.62	0.48	4.78	21.08	27.29	38.48	61.95
C12CM_spot	-0.03	-0.12	-0.32	-0.99	-0.66	0.65	-1.06
	(-1.12)	(*-1.73*)	(-1.04)	(-1.13)	(-1.02)	(0.96)	(**-2.96**)
R^2 (%)	0.09	2.42	3.41	9.90	3.25	2.06	7.50
CVAR	-13.98	-26.18	-0.98	-151.47	-198.14	-197.79	-392.73
	(**-2.95**)	(**-2.68**)	(-0.05)	(-0.88)	(**-2.85**)	(**-3.10**)	(-1.59)
R^2 (%)	4.24	3.56	-0.89	0.73	1.52	1.37	3.62
CCM	-0.01	-0.03	-0.17	-0.36	-0.45	-0.58	-0.84
	(-0.50)	(-0.70)	(-1.11)	(-2.43)	(**-4.55**)	(**-5.39**)	(**-17.51**)
R^2 (%)	-0.59	0.59	8.72	29.95	40.37	47.52	79.03
C12CM	-0.02	-0.08	-0.31	-0.51	-0.61	-0.65	-1.21
	(-0.81)	(-1.61)	(*-1.81*)	(-1.25)	(-1.60)	(-2.57)	(**-6.10**)
R^2 (%)	-0.34	2.15	7.23	4.49	6.72	8.34	33.08
CVAL_CM	0.01	0.01	-0.06	-0.27	-0.59	-0.48	0.13
	(0.97)	(0.26)	(-0.53)	(-1.10)	(**-5.68**)	(**-3.26**)	(0.50)
R^2 (%)	-0.41	-0.79	-0.53	3.80	22.59	15.57	-0.59
CMOM_CM	-0.04	-0.12	-0.31	-0.14	0.54	0.59	0.48
	(-2.44)	(-2.39)	(-2.06)	(-1.07)	(**2.79**)	(**4.10**)	(2.47)
R^2 (%)	3.48	8.03	10.74	0.67	25.38	29.44	5.92

Regarding the first subsample, from January 1972 to December 1999 (Panel B of Table 10), the regression coefficients of *D/E, SVAR, B/M, CSP, DFR, I/K, IP, GDP,* and *CCM_spot* are insignificant at the 5% level for each horizon. Hence, the significance of *D/E* and *DFR* vanishes when I consider only the first subsample. On the other hand, *LTR* now significantly forecasts positive returns at the 60-month horizon (according to the Hansen–Hodrick t-ratio), while *DEF* and *INFL* become significant predictors of negative returns at several horizons.

Moreover, the results for *C12CM_spot* at the one- to 12-month horizons indicate time-series momentum within this subsample: A high cumulative spot return over the last 12 months forecasts high future short-term returns, which coincides with the findings of Moskowitz et al. (2012). The predictors introduced in the first subsample are *CSP* and *CP*. While the slope of *CSP* is insignificant, *CP* significantly predicts negative returns at horizons of one month and three months (at the 10% level) and positive returns at the 48-month horizon (according to the Hansen–Hodrick t-ratio). Moreover, *DY, E/P*, *CRMRF*, *CL*, *SENT*, *RF*, *LTY*, and *CAY* still consistently forecast negative returns, while *TERM* and *M2* still predict positive returns.

Examining the second subsample, from January 2000 to June 2010 (Panel C of Table 10), one can see that, in contrast to the first subsample, all candidate predictors, except *DY*, are now significant for at least one horizon. Consequently, predictability appears to have increased at the millennium. Moreover, there are several variables whose predictive signs have changed compared to the first subsample (at least for some horizons). For instance, *DEF* and *CAY* now predict long-horizon returns significantly positively. Thus, the results indicate some degree of structural change in commodity markets over time.

Table 11 displays the results of the IS model selection procedure. First, observe that the procedure does not select the specification $n = 0$ for either forecasting horizon. Instead, seven predictors are chosen for each horizon. Consequently, these results confirm the suggestion from the single regressions that one can reject the null hypothesis of returns that are unpredictable beyond a constant, that is, returns that are i.i.d. through time. Second, observe that the adjusted R^2 value again increases with horizons and is considerably higher than for the single regressions. Third, note that the variables that are chosen by the model selection procedure and which thus build the best predictive model depend on both the horizon and the sample period. Nevertheless, some predictors seem to be particularly important and robust, being represented in most models: *CCM_spot* (19 out of 21 models), *M2* (15 models), *CAY* (14 models), and *SENT* (13 models).[36] Hence, a combination of the current spot time-series value, monetary policy, consumption–wealth ratio, and investor sentiment seems to represent a good portion of return predictability. The variable *INFL* is the sole variable that is not represented in any of the best models.

[36] It is remarkable that *CCM_spot* is insignificant within all single predictive regressions.

Table 11. Multiple predictive regressions.

This table displays the results for multiple-variable predictive regressions for the monthly continuously compounded return on an equal-weighted portfolio of 27 commodity futures at horizons q = 1, 3, 12, 24, 36, 48, and 60 months. The returns do not include the returns on collateral from transacting in futures contracts. The multiple-variable models are chosen by a model selection procedure that searches for the model with the maximum adjusted R^2 out of all possible combinations of the variables shown in the table, with a maximum of seven variables at once. The original sample period is January 1972 to June 2010 (Panel A), the two subsample periods are January 1972 to December 1999 (Panel B) and January 2000 to June 2010 (Panel C), respectively, and q observations are lost in each of the respective q-horizon regressions. The first line corresponding to each model reports the slope estimates. A hyphen for the slope estimate means that the selection procedure does not include this variable in the model. Line 2 reports the Newey–West t-ratios (in parentheses) and line 3 shows the Hansen–Hodrick t-ratios (in square brackets), both computed with q lags. Italic, underlined, and boldfaced t-statistics denote statistical significance according to the standard normal distribution at the 10%, 5%, and 1% levels, respectively. The last line, R^2 (%), shows the values of the adjusted coefficient of determination (%).

Panel A: January 1972 to June 2010

	DY	*E/P*	*D/E*	*SVAR*	*B/M*	*CRMRF*	*CL*	*SENT*	*RF*	*LTY*	*LTR*	*TERM*	*DEF*	*DFR*	*INFL*	*I/K*	*CAY*	*IP*	*M2*	*GDP*	*CCM_spot*	*C12CM_spot*	*CVAR*	R^2 (%)
q = 1	–	–	0.01	–	–	–	–	-0.01	–	–	–	–	–	0.33	–	–	-0.43	–	–	–	-0.04	0.05	-8.32	7.85
			(1.30)					(**-3.90**)						(1.24)			(**-4.27**)				(**-3.41**)	(2.36)	(-2.45)	
			[1.24]					[**-3.80**]						[1.24]			[**-4.13**]				[**-3.30**]	[2.32]	[-2.38]	
q = 3	0.10	–	–	–	-0.13	–	–	-0.04	–	–	–	–	–	–	–	–	-2.05	–	–	–	-0.16	0.11	-16.63	17.01
	(2.30)				(*-1.78*)			(**-4.39**)									(**-3.94**)				(**-3.59**)	(2.34)	(-2.48)	
	[2.00]				[-1.56]			[**-3.76**]									[**-3.33**]				[**-3.05**]	[2.12]	[-2.37]	
q = 12	–	–	–	–	–	-0.30	0.23	-0.17	–	–	–	–	–	–	–	–	-4.08	–	0.77	–	-0.54	0.32	–	47.40
						(**-4.21**)	(**3.16**)	(**-5.77**)									(**-3.90**)		(2.27)		(**-4.85**)	(2.37)		
						[**-3.76**]	[**2.72**]	[**-5.09**]									[**-3.59**]		[*1.88*]		[**-4.53**]	[2.02]		
q = 24	–	–	–	–	–	-0.64	0.37	-0.21	–	-2.95	–	–	–	–	–	–	-3.95	–	1.71	–	-0.73	–	–	61.59
						(**-4.45**)	(**2.86**)	(**-3.62**)		(*-1.66*)							(-1.64)		(**2.69**)		(**-4.56**)			
						[**-3.91**]	[**2.60**]	[**-3.21**]		[-1.64]							[-1.50]		[2.52]		[**-4.44**]			
q = 36	–	–	0.41	–	–	-0.37	–	–	-6.36	–	–	–	–	–	–	–	-2.87	0.99	1.23	–	-0.32	–	–	65.35
			(1.63)			(**-2.79**)			(**-7.37**)								(-1.12)	(2.44)	(2.41)		(-2.44)			
			[1.47]			[**-2.66**]			[**-6.50**]								[-1.04]	[**2.62**]	[**2.85**]		[-2.09]			
q = 48	–	–	0.45	–	–	–	–	-0.08	-4.98	–	–	–	–	–	–	–	-4.75	–	2.23	-1.63	-0.32	–	–	64.61
			(*1.80*)					(*-1.83*)	(**-2.60**)								(-1.48)		(**4.76**)	(-0.93)	(*-1.81*)			
			[*1.66*]					[**-3.32**]	[-2.43]								[-1.46]		[**10.14**]	[-0.83]	[*-1.73*]			
q = 60	–	-0.67	–	–	–	-0.48	0.16	-0.13	–	–	–	–	11.17	–	–	–	–	–	2.87	–	-0.33	–	–	62.40
		(**-3.42**)				(*-1.81*)	(0.83)	(-2.36)					(0.86)						(**3.11**)		(-1.18)			
		[**-2.85**]				[-1.58]	[0.79]	[**-2.71**]					[0.88]						[**3.84**]		[-1.00]			

Table 11 (continued)

Panel B: January 1972 to December 1999

	DY	*E/P*	*D/E*	*SVAR*	*B/M*	*CRMRF*	*CL*	*SENT*	*RF*	*LTY*	*LTR*	*TERM*	*DEF*	*DFR*	*INFL*	*I/K*	*CAY*	*IP*	*M2*	*GDP*	*CCM_spot*	*C12CM_spot*	*CVAR*	R^2 (%)
$q = 1$	0.03 (2.52) [2.45]	–	–	–	–	–	–	0.01 (2.31) [2.30]	–	-1.25 (**-4.02**) [**-4.04**]	-0.13 (*-1.76*) [*-1.74*]	–	–	–	–	–	–	–	0.10 (**2.94**) [**2.89**]	–	–	0.04 (1.43) [1.43]	-7.13 (-1.38) [-1.33]	10.15
$q = 3$	–	0.04 (1.36) [1.23]	–	–	–	–	–	–	–	-1.92 (**-4.49**) [**-3.95**]	–	–	–	–	–	-5.29 (-2.47) [-2.11]	-1.10 (*-1.90*) [-1.63]	–	0.18 (*1.76*) [1.59]	–	-0.13 (**-3.26**) [**-2.96**]	0.17 (**3.23**) [**3.50**]	–	24.58
$q = 12$	–	–	-0.30 (*-1.94*) [*-1.86*]	–	–	-0.27 (**-2.61**) [-2.32]	–	–	–	-5.39 (**-5.37**) [**-4.84**]	–	–	–	–	–	-29.89 (**-3.03**) [**-3.08**]	-7.77 (**-4.47**) [**-4.12**]	–	–	–	-0.77 (**-5.59**) [**-5.71**]	0.44 (**2.94**) [**2.63**]	–	59.59
$q = 24$	–	–	–	–	-0.59 (**-3.26**) [**-2.86**]	-0.56 (**-3.53**) [**-3.38**]	–	-0.16 (**-5.29**) [**-4.41**]	–	–	–	–	–	–	–	–	-13.35 (**-5.36**) [**-5.20**]	–	1.14 (1.96) [*1.88*]	-1.73 (*-1.89*) [-1.56]	-1.22 (**-6.65**) [**-6.04**]	–	–	73.57
$q = 36$	–	–	–	–	-1.21 (**-5.44**) [**-4.60**]	-0.70 (**-4.06**) [**-3.67**]	–	–	-4.16 (**-3.13**) [-2.57]	–	–	–	–	–	–	–	-11.72 (**-5.86**) [**-6.10**]	–	1.76 (**4.27**) [**4.83**]	–	-0.60 (**-2.92**) [-2.54]	-0.60 (-2.32) [*-1.83*]	–	79.67
$q = 48$	-2.34 (**-7.86**) [**-8.38**]	–	–	–	0.79 (2.49) [**2.62**]	-0.80 (**-9.96**) [**-8.91**]	–	-0.10 (**-2.61**) [**-2.81**]	–	–	–	–	28.96 (**3.30**) [**2.80**]	–	–	–	–	–	2.73 (**10.62**) [**11.28**]	–	-1.00 (**-5.10**) [**-5.42**]	–	–	82.71
$q = 60$	-2.54 (**-6.09**) [**-5.41**]	–	–	–	1.21 (2.40) [2.03]	-0.99 (**-8.12**) [**-8.78**]	–	–	–	–	–	–	23.87 (2.45) [**2.85**]	–	–	–	–	–	2.93 (**3.89**) [**3.75**]	-2.15 (**-3.45**) [**-5.75**]	-0.83 (**-5.25**) [**-4.63**]	–	–	80.64

Table 11 (continued)

Panel C: January 2000 to June 2010

	DY	*E/P*	*D/E*	*SVAR*	*B/M*	*CRMRF*	*CL*	*SENT*	*RF*	*LTY*	*LTR*	*TERM*	*DEF*	*DFR*	*INFL*	*I/K*	*CAY*	*IP*	*M2*	*GDP*	*CCM_spot*	*C12CM_spot*	*CVAR*	R^2 (%)
q = 1	–	–	–	–	-0.20	–	–	-0.04	–	-2.18	–	–	–	0.32	–	–	-2.01	–	–	–	-0.11	–	-9.49	26.19
					(**-2.89**)			(**-5.59**)		(-2.40)				(1.05)			(**-3.14**)				(**-2.92**)		(**-2.65**)	
					[**-3.60**]			[**-5.55**]		[-2.23]				[1.08]			[**-2.87**]				[**-2.68**]		[**-3.50**]	
q = 3	–	–	–	–	–	–	–	-0.04	–	-4.75	–	–	–	–	–	-16.70	-8.45	–	2.17	1.09	-0.37	–	–	46.43
								(**-3.07**)		(-1.99)						(-2.14)	(**-3.31**)		(**2.80**)	(2.49)	(**-2.82**)			
								[**-3.01**]		[-2.21]						[*-1.84*]	[**-2.97**]		[**2.70**]	[2.26]	[**-2.70**]			
q = 12	–	0.19	–	6.28	–	–	–	–	8.15	–	–	–	–	–	–	-94.97	-9.36	–	2.81	–	-0.80	–	–	72.82
		(**5.19**)		(**4.19**)					(2.54)							(**-6.69**)	(**-4.49**)		(2.44)		(**-7.06**)			
		[**14.53**]		[**3.47**]					[2.07]							[**-5.97**]	[**-4.36**]		[2.46]		[**-8.81**]			
q = 24	–	–	–	–	–	–	–	-0.02	–	9.21	–	–	9.86	–	–	–	–	-3.92	-3.92	–	-1.03	0.46	–	91.45
								(-1.63)		(**2.58**)			(*1.97*)					(**-19.58**)	(**-2.69**)		(**-7.08**)	(1.61)		
								[*-1.83*]		[2.34]			[*1.78*]					[**-40.78**]	[-2.39]		[**-5.95**]	[1.45]		
q = 36	–	–	0.69	–	–	–	-0.37	–	-10.81	-8.84	–	–	–	–	–	65.98	–	–	–	–	0.39	-0.70	–	89.33
			(2.45)				(**-3.15**)		(**-5.51**)	(-1.97)						(2.04)					(1.19)	(**-3.80**)		
			[2.15]				[**-2.75**]		[**-8.17**]	[-1.97]						[*1.93*]					[1.04]	[**-4.39**]		
q = 48	–	–	–	–	–	–	0.28	0.11	–	–	–	5.60	23.79	–	–	–	–	1.23	–	-7.66	–	0.68	–	92.09
							(**3.77**)	(**7.94**)				(**3.59**)	(**5.35**)					(**8.04**)		(**-7.99**)		(2.05)		
							[**4.18**]	[**45.23**]				[**-3.40**]	[**11.71**]					[**19.96**]		[**8.21**]		[**7.65**]		
q = 60	–	–	0.68	–	–	-0.84	–	–	11.80	–	–	–	-17.83	–	–	–	-11.98	–	8.71	–	-1.00	–	–	93.70
			(**9.03**)			(**-5.84**)			(**6.04**)				(**-2.82**)				(**-6.31**)		(**8.13**)		(**-7.22**)			
			[**16.42**]			[**22.47**]			[**-27.74**]				[**13.04**]				[**-83.46**]		[**-50.14**]		[**-24.45**]			

5.1.2. Out-of-sample results

Panel A of Table 12 displays the out-of-sample (OOS) forecasting results over the OOS period January 1980 to June 2010. We see that the combining methods proposed by Rapach et al. (2010) perform quite well, with R^2_{OS} between 0.96% (median combination) and 1.65% (mean combination). Moreover, the p-values both from the bootstrap procedure and obtained using Clark and West's (2007) *MSPE-adjusted* statistic indicate statistical significance at the 1% level.

Table 12. OOS forecasting.

This table reports the results of OOS forecasts of the monthly return on an equal-weighted portfolio of 27 commodity futures. The returns do not include the returns on collateral from transacting in futures contracts. The forecasts are obtained from individual predictive regressions, from mean, median, and trimmed mean combination methods, as well as from a model selection procedure. The OOS evaluation periods are January 1980 to June 2010 (Panel A), January 1980 to December 1999 (Panel B), and January 2000 to June 2010 (Panel C). The IS estimation periods start eight years prior to the OOS periods. The variable R^2_{OS} (%) denotes Campbell and Thompson's (2008) OOS R^2 statistic (%) and $HM-S$ denotes the Henriksson–Merton (1981) test statistic. Italics, an underline, and boldface indicate statistical significance according to a bootstrap procedure at the 10%, 5%, and 1% levels, respectively. The term p(CW) is the p-value for rejecting the null hypothesis $R^2_{OS} \leq 0$ according to the *MSPE-adjusted* statistic of Clark and West (2007).

Comb. method or predictor	R^2_{OS} (%)	p(CW)	$HM-S$	Predictor	R^2_{OS} (%)	p(CW)	$HM-S$
Panel A: OOS period January 1980 to June 2010							
Mean	**1.65**	0.002	1.00	*LTR*	-0.42	0.438	0.99
Median	**0.96**	0.000	0.99	*TERM*	-2.56	0.918	1.00
Trimmed mean	**1.37**	0.002	1.00	*DEF*	-1.98	0.021	1.06
Model selection	-7.36	0.018	1.06	*DFR*	*0.30*	0.210	0.97
DY	-2.72	0.075	1.02	*INFL*	-1.13	0.510	0.99
E/P	-2.22	0.207	1.02	*I/K*	-2.93	0.922	1.00
D/E	-1.39	0.615	1.01	*CAY*	-3.63	0.005	0.98
SVAR	-4.66	0.691	1.01	*IP*	-5.19	0.738	0.99
B/M	-2.58	0.256	*1.03*	*M2*	-1.34	0.002	1.01
CRMRF	-0.71	0.279	0.98	*GDP*	-0.57	0.477	0.99
CL	-2.97	0.737	0.97	*CCM_spot*	-0.95	0.666	0.97
SENT	-2.78	0.450	**1.09**	*C12CM_spot*	-0.27	0.035	1.08
RF	-0.33	0.042	1.07	*CVAR*	*0.28*	0.218	1.01
LTY	-0.21	0.001	1.06				
Panel B: OOS period January 1980 to December 1999							
Mean	**2.88**	0.002	1.00	*LTR*	-0.37	0.415	1.00
Median	**1.83**	0.000	0.99	*TERM*	-4.56	0.908	1.00
Trimmed mean	**2.52**	0.002	1.00	*DEF*	-3.08	0.019	*1.04*
Model selection	-9.77	0.044	1.04	*DFR*	-0.36	0.489	0.98
DY	-4.91	0.071	1.02	*INFL*	-1.61	0.441	0.98
E/P	-3.82	0.182	1.03	*I/K*	-5.13	0.905	0.98
D/E	-1.09	0.603	1.00	*CAY*	-6.31	0.007	0.89
SVAR	-8.37	0.757	0.99	*IP*	-8.89	0.701	0.98
B/M	-4.61	0.240	*1.04*	*M2*	-2.58	0.002	0.95
CRMRF	-1.34	0.357	0.96	*GDP*	-0.34	0.332	1.00
CL	-6.01	0.793	0.93	*CCM_spot*	-0.98	0.550	0.95
SENT	-8.18	0.729	1.03	*C12CM_spot*	**4.27**	0.000	1.08
RF	1.50	0.025	1.10	*CVAR*	0.02	0.400	1.00
LTY	**2.72**	0.001	*1.06*				

Table 12 (continued)

Comb. method or predictor	R^2_{OS} (%)	p(CW)	$HM-S$	Predictor	R^2_{OS} (%)	p(CW)	$HM-S$
Panel C: OOS period January 2000 to June 2010							
Mean	1.05	0.149	1.08	*LTR*	-1.25	0.686	1.09
Median	-0.24	0.690	1.06	*TERM*	-0.84	0.584	1.03
Trimmed mean	0.37	0.233	1.08	*DEF*	-4.61	0.603	1.04
Model selection	-19.85	0.037	1.07	*DFR*	-2.26	0.269	0.95
DY	-0.89	0.884	**1.17**	*INFL*	-2.38	0.657	1.03
E/P	-2.35	0.675	1.14	*I/K*	*0.58*	0.174	1.11
D/E	-2.60	0.488	1.10	*CAY*	-1.23	0.824	1.14
SVAR	-7.15	0.175	1.09	*IP*	-0.59	0.399	*1.07*
B/M	-0.53	0.833	*1.07*	*M2*	-1.17	0.862	1.11
CRMRF	-0.10	0.281	*1.07*	*GDP*	-2.42	0.771	*1.08*
CL	0.08	0.266	1.01	*CCM_spot*	-1.85	0.807	0.93
SENT	**3.80**	0.001	*1.08*	*C12CM_spot*	-1.68	0.681	1.10
RF	-1.47	0.654	1.06	*CVAR*	-23.63	0.094	1.04
LTY	-2.19	0.828	1.10				

According to Campbell and Thompson (2008) and Rapach et al. (2010), even a small positive R^2_{OS}, such as the 0.5% for a monthly sample frequency, can indicate a degree of return predictability that is economically meaningful. Thus, the R^2_{OS} value I obtain from the combination methods are economically significant and indicate that aggregate returns on commodity futures are predictable OOS. The R^2_{OS} values of the model selection procedure as well as of the individual forecasts paint a picture that is less favorable for the alternative hypothesis of return predictability. First, though the Clark–West (2007) p-values that I obtain for the model selection procedure indicate that the null hypothesis of having a higher mean-squared prediction error than the historical average forecast can be rejected at the 5% level, neither the point estimate of the R^2_{OS} or its p-value obtained from the bootstrap procedure support this. Second, all but two of the individual forecasts' R^2_{OS} are negative. Some individual forecasts show Clark–West (2007) p-values that indicate that the null hypothesis of having a higher mean-squared prediction error than the historical average forecast can be rejected at the 10% level and below (*DY*, *RF*, *LTY*, *DEF*, *CAY*, *M2*, and *C12CM_spot*). However, both the R^2_{OS} point estimates and the bootstrap p-values argue against this supposition.

The R^2_{OS} statistic measures the closeness of forecasted returns to actually realized returns. Thus, a forecaster who is concerned about the ability of the forecasting model to correctly predict the magnitude of future returns should regard the R^2_{OS} and choose a combination method according to my results. However, forecasters who are more interested in a forecasting model's ability to provide the correct sign of future returns (e.g., an investor who is only interested in the correct *direction* for trading activity) should regard the Henriksson–Merton test statistic (Zakamulin, 2013). According to the values that I obtain for this measure, the model selection procedure as well as individual forecasts from single predictive regressions, including *SENT*, *RF*, *LTY*, *DEF*, and *C12CM_spot*, seem to outperform the historical mean model, with $HM-S$ values between 1.06 and 1.09 that are statistically significant at the 5% level and below. In contrast, the combining methods do not seem to outperform the historical average with regard to $HM - S$.

Panel B of Table 12 shows the results for the OOS period January 1980 to December 1999. Basically, the combination methods perform better, to some extent, than during January 1980 to June 2010, with R_{OS}^2 values between 1.83% (median combination) and 2.88% (mean combination). Moreover, the model selection procedure still performs poorly in terms of R_{OS}^2. Furthermore, two individual forecasts show a performance that is comparable to the combined forecasts: RF and LTY, with R_{OS}^2 values of 1.50% and 2.72%, respectively, statistically significant at the 5% and 1% levels, respectively, according to both the p-values from bootstrapping and from Clark and West (2007). However, the best performance is that of the forecast from a predictive regression with *C12CM_spot* as the single predictor variable (R_{OS}^2 is 4.27% and statistically significant at the 1% level according to both bootstrap and Clark and West (2007) p-values). Significantly positive $HM - S$ values are shown by the individual forecasts from these three variables, as well as from *B/M* and *DEF*. Overall, the predictability of the magnitude of future returns as measured by R_{OS}^2 appears to be somewhat higher within the January 1980 to December 1999 period than in the January 1980 to June 2010 period.

Finally, the results for the OOS period January 2000 to June 2010 are displayed in Panel C of Table 12. On the one hand, we see that the R_{OS}^2 values of the three combination methods are much lower than for the two other OOS periods. Only the R_{OS}^2 value of the mean combination method is significantly positive and only according to the p-value obtained from the bootstrap. Furthermore, the model selection procedure performs very poorly according to R_{OS}^2. Moreover, only $SENT$ generates an individual forecast whose R_{OS}^2 of 3.80% is significantly positive according to both methods for p-value estimation. On the other hand, the majority of forecasting models (including the three combining methods, but not the model selection procedure) generates significantly positive $HM - S$ values. Thus, the direction of future returns appears to be more predictable within the period January 2000 to June 2010 than within the two other OOS periods considered.

Overall, the results indicate that both the magnitudes and signs of future aggregate commodity returns are predictable OOS. However, predictability with regard to the correct forecast magnitude relies mainly on the application of forecast combination methods, whereas predictability with regard to the correct forecast direction is mainly based on individual predictive regressions.

5.2. Results for Expected Returns in the Cross Section of European Stocks

In the following sections, I present the evidence on multifactor models and their consistency with the Intertemporal CAPM (ICAPM) that I obtain from the European stock market (Research Questions 2 and 3).[37]

[37] This chapter is based on Secs. 5 and 6 of Lutzenberger (2014b).

5.2.1. Results from predictive regressions

Table 13 displays the results for the multiple long-horizon regressions conducted with $TERM$, DEF, DY, RF, PE, VS, and CP as the right-hand side variables. The table presents only the results for the one-, 12-, and 60-month horizons, while the results for the remaining horizons (three months, 24 months, 36 months, and 48 months) are available on request. For $q = 1$ and $q = 12$ (Panels A and B), $TERM$ forecasts significantly positive market returns within the regressions corresponding to the ICAPM of Hahn and Lee (2006), HL; the unrestricted version of the ICAPM of Campbell and Vuolteenaho (2004), CV; and the three-factor model of Koijen et al. (2010), KLVN. Regarding Panel C ($q = 60$), the predictive slope of $TERM$ is only significant within the ICAPM of Petkova (2006), P, where it also shows a positive sign. The slopes of DEF are negative for $q = 1$ and $q = 60$ within HL (but without significance) and negative for all three horizons within P (significant at the 10% level for $q = 12$). In contrast, DY significantly forecasts positive market returns at horizons of 12 months and 60 months. Moreover, RF forecasts returns at the 60-month horizon to be significantly positive, while the 60-month forecast of PE is significantly negative. Finally, the predictive slope of VS is positive but insignificant for all three horizons, while that for CP is significant and negative for $q = 1$ and $q = 60$.

To determine whether ICAPM criteria 2, 3a, and 3b are satisfied, I make an overall assessment of the significance and signs of the predictive slopes for each state variable within each model in view of my results for all seven forecasting horizons. In this overall assessment, I rate a state variable as a significant predictor if its predictive slope is significant at least at the 10% level for at least one forecasting horizon. To make an overall assessment of the predictive slopes' signs, I consider only coefficients that are significant. If none of the slopes is significant, I also consider the insignificant coefficients' signs. If the signs are inconsistent across (significant) forecasting horizons, I assess the overall sign as *indeterminate* and later interpret the consistency with the risk price estimates according to both possible signs (i.e., positive and negative). The results are presented in Panel D of Table 13.

Altogether, $TERM$ significantly forecasts positive market returns within all four models. In contrast, the sign of the slope of DEF is indeterminate and insignificant within HL, countering the US results, where DEF forecasts positive market returns conditional on $TERM$. Within P, the slope of DEF is significantly negative. The variables DY and RF both predict positive market returns within P, whereas PE forecasts negative returns within CV. Finally, the slope of VS is positive (but nonsignificant), while CP forecasts negative returns. The latter results do not agree with the evidence of Maio and Santa-Clara (2012, MSC) or the original study of Cochrane and Piazzesi (2005), where CP significantly forecasts positive US market returns.

Table 13. Multiple predictive regressions for ICAPM state variables.

This table shows the results for the multiple predictive regressions with the monthly continuously compounded return of the European stock market as the left-hand side variable for horizons of one month (Panel A), 12 months (Panel B), and 60 months (Panel C). The right-hand side variables are the current European values of the slope of the yield curve $TERM$, the corporate bond default spread DEF, the market dividend-to-price ratio DY, the short-term risk-free rate RF, the aggregate price–earnings ratio PE, the value spread VS, and the Cochrane–Piazzesi factor CP. The original sample period is December 1990 to December 2011. However, one, 12, and 60 observations are lost in each of the one-, 12-, and 60-month horizon regressions, respectively. The first row of each regression shows the estimated slope coefficients and the second row displays Newey–West t-ratios calculated with one, 12, or 60 lags (in parentheses). The level of statistical significance of the estimated slopes is indicated by italics (10% level), an underline (5% level), and boldface (1% level). The R^2 (%) value presents the *adjusted* R^2 (%) value. Panel D shows the overall assessment of the variables' predictive characteristics, jointly considering the regression results for forecasting horizons of one month, three months, 12 months, 24 months, 36 months, 48 months, and 60 months. A + sign (- sign) means that the variable all in all forecasts positive (negative) market returns, while +/- indicates that the sign of the respective predictive slope is assessed as indeterminate. Parentheses indicate that the forecast is insignificant for all forecasting horizons.

Row	TERM	DEF	DY	RF	PE	VS	CP	R^2 (%)
Panel A: q = 1								
1	0.41	-0.71						1.59
	(1.99)	(-0.67)						
2	0.39	-1.95	0.01	-0.10				2.06
	(1.07)	(-1.43)	(1.59)	(-0.40)				
3	0.45				0.00	0.01		0.80
	(2.16)				(0.05)	(0.49)		
4	0.80						-0.83	2.42
	(**2.75**)						(*-1.85*)	
Panel B: q = 12								
1	5.13	0.33						11.44
	(*1.75*)	(0.05)						
2	4.34	-23.36	0.20	-2.05				30.95
	(1.13)	(*-1.94*)	(**2.76**)	(-0.91)				
3	5.22				-0.10	0.10		13.82
	(*1.88*)				(-1.15)	(0.45)		
4	6.82						-4.02	12.92
	(2.33)						(-1.29)	
Panel C: q = 60								
1	-0.73	-25.52						0.45
	(-0.12)	(-0.48)						
2	20.96	-23.55	0.40	11.23				62.86
	(**3.57**)	(-0.50)	(**4.56**)	(2.04)				
3	5.27				-1.01	0.83		27.52
	(0.79)				(**-2.75**)	(1.10)		
4	5.71						-13.55	3.43
	(0.78)						(-2.12)	
Panel D: Overall assessment								
1	+	(+/-)						
2	+	-	+	+				
3	+				-	(+)		
4	+						-	

Table 14 displays the results for the predictive regressions with state variables constructed from empirical factors. The slope of SMB^* is insignificant for all three horizons displayed in the table and for all models. In contrast, the slope of HML^* is significantly positive for $q =$

12—within the four-factor model of Carhart (1997), C; the four-factor model of Pástor and Stambaugh (2003), PS; and the five-factor model of Fama and French (1993), FF5—and for $q = 60$ (within all models). Moreover, the slope of $CUMD$ is significantly positive for $q = 12$ and $q = 60$ and that of CL is significantly positive for $q = 12$. Finally, the slope of $TERM$ is significantly positive for $q = 1$ and $q = 12$, whereas that of DEF is negative but insignificant for all three horizons.

My overall assessment with regard to all seven forecasting horizons is displayed in Panel D of Table 14. Overall, the predictive slope on SMB^* is positive but insignificant within the three-factor model of Fama and French (1993), FF3. Within C, PS, and FF5, the slope of SMB^* is insignificant except for the 36-month horizon (available upon request), where it is significantly negative. Consequently, the slope is significant and negative in my overall assessment. This result is at odds with the US evidence, where SMB^* forecasts significantly positive returns within all models. A potential explanation for this differing result might be the negative size effect I detect in my European sample. Moreover, the slope of HML^* is significantly positive within all four models. The variables $CUMD$, CL, and $TERM$ also significantly forecast positive returns. Finally, the slope of DEF is negative but insignificant, contrary to the US evidence.

Next, I regard the results for the multiple long-horizon regressions with $SVAR$ as the left-hand side variable and ICAPM state variables on the right-hand side, as displayed in Table 15. For $q = 1$, the slope of $TERM$ is significantly negative within P, while it is significant within all four models for the 60-month horizon (positive within HL, CV, and KLVN, but negative within P). Furthermore, the slope of DEF is significantly positive for the one- and 12-month horizons within both HL and P. For $q = 60$, the slope is significantly negative within P. In addition, DY forecasts market volatility at the 12- and 60-month horizons significantly negatively, as RF does for $q = 1$ and $q = 60$. The variable PE behaves inconsistently across horizons: The one-month forecast is significantly negative, whereas the 60-month forecast is significantly positive. The variable VS behaves similarly, but with opposite signs: The variable forecasts market volatility at $q = 1$ significantly positively, whereas its slope for $q = 60$ is significantly negative. Finally, the slope of CP is positive but insignificant for all three horizons.

I present the results of my overall assessment regarding all seven forecasting horizons in Panel D of Table 15. All in all, $TERM$ significantly forecasts increasing market volatility within HL, CV, and KLVN, whereas its predictive slope is significantly negative within P. The variable DEF significantly forecasts increasing volatility within HL. In contrast, the slope of DEF within P is assessed as indeterminate: The slope is significantly positive for the one-, three-, 12-, 24-, and 36-month horizons, but significantly negative for the 60-month horizon. The variables DY and RF predict decreasing volatility, whereas the slope of PE is indeterminate: The one- and three-month forecasts are significantly negative, whereas the 36-, 48-, and 60-month forecasts are significantly positive. The same applies to VS. Finally, the predictive slope of CP is positive but insignificant.

Table 14. Multiple predictive regressions for state variables constructed from empirical factors.

This table shows the results for the multiple predictive regressions with the monthly continuously compounded return of the European stock market as the left-hand side variable for horizons of one month (Panel A), 12 months (Panel B), and 60 months (Panel C). The right-hand side variables are the current European values of the size premium SMB^{*}, the value premium HML^{*}, the momentum premium $CUMD$, the liquidity factor CL, the slope of the yield curve $TERM$, and the corporate bond default spread DEF. The original sample period is December 1990 to December 2011. However, one, 12, and 60 observations are lost in each of the one-, 12-, and 60-month horizon regressions and the original sample period for the regressions containing $CUMD$ or CL is November 1995 to December 2011. The first row of each regression shows the estimated slope coefficients and the second row displays Newey–West t-ratios calculated with one, 12, or 60 lags (in parentheses). The levels of statistical significance of the estimated slopes are indicated by italics (10% level), an underline (5% level), and boldface (1% level). The R^2 (%) value presents the *adjusted* R^2 value (%). Panel D shows the overall assessment of the variables' predictive characteristics, jointly considering the regression results for forecasting horizons of one month, three months, 12 months, 24 months, 36 months, 48 months, and 60 months. A + sign (- sign) means that the variable all in all forecasts positive (negative) market returns, while +/- indicates that the sign of the respective predictive slope is assessed as indeterminate. Parentheses indicate that the forecast is insignificant for all forecasting horizons.

Row	SMB*	HML*	CUMD	CL	TERM	DEF	R² (%)
Panel A: q = 1							
1	0.01	-0.00					0.06
	(1.23)	(-0.39)					
2	0.01	-0.00	0.00				-0.55
	(1.31)	(-0.30)	(0.28)				
3	0.01	-0.00		0.38			0.54
	(1.41)	(-0.12)		(1.50)			
4	0.01	0.00			0.40	-0.69	1.37
	(0.83)	(0.14)			(*1.92*)	(-0.64)	
Panel B: q = 12							
1	0.00	0.03					3.12
	(0.04)	(1.52)					
2	0.02	0.05	0.40				9.03
	(0.16)	(2.41)	(2.30)				
3	0.02	0.04		4.66			7.46
	(0.13)	(2.04)		(*1.87*)			
4	-0.02	0.05			5.27	-3.12	16.09
	(-0.20)	(2.44)			(1.97)	(-0.41)	
Panel C: q = 60							
1	0.04	0.20					36.44
	(0.22)	(**4.72**)					
2	-0.21	0.13	0.43				31.71
	(-1.26)	(**5.86**)	(*1.88*)				
3	-0.18	0.12		-6.59			28.71
	(-0.96)	(**4.93**)		(-0.77)			
4	-0.08	0.23			0.85	-46.60	40.12
	(-0.45)	(**3.69**)			(0.23)	(-1.10)	
Panel D: Overall assessment							
1	(+)	+					
2	-	+	+				
3	-	+		+			
4	-	+			+	(-)	

Table 15. Multiple predictive regressions for ICAPM state variables ($SVAR$).

This table shows the results for our multiple predictive regressions with the variance ($SVAR$) of the European stock market as the left-hand side variable for horizons of one month (Panel A), 12 months (Panel B), and 60 months (Panel C). The right-hand side variables are the current European values of the slope of the yield curve $TERM$, the corporate bond default spread DEF, the market dividend-to-price ratio DY, the short-term risk-free rate RF, the aggregate price–earnings ratio PE, the value spread VS, and the Cochrane–Piazzesi factor CP. The original sample period is December 1990 to December 2011. However, one, 12, and 60 observations are lost in each of the one-, 12-, and 60-month horizon regressions. The first row of each regression shows the estimated slope coefficients and the second row displays the Newey–West t-ratios calculated with one, 12, or 60 lags (in parentheses). The levels of statistical significance of the estimated slopes are indicated by italics (10% level), an underline (5% level), and boldface (1% level). The R^2 (%) value presents the *adjusted* R^2 (%) value. Panel D shows the overall assessment of the variables' predictive characteristics, jointly considering the regression results for forecasting horizons of one month, three months, 12 months, 24 months, 36 months, 48 months, and 60 months. A + sign (- sign) means that the variable all in all forecasts positive (negative) market returns, while +/- indicates that the sign of the respective predictive slope is assessed as indeterminate. Parentheses indicate that the forecast is insignificant for all forecasting horizons.

Row	TERM	DEF	DY	RF	PE	VS	CP	R² (%)
Panel A: q = 1								
1	-0.02	0.35						17.41
	(-0.76)	(**3.56**)						
2	-0.06	0.31	-0.00	-0.03				18.56
	(*-1.88*)	(**2.92**)	(-0.30)	(*-1.81*)				
3	-0.03				-0.00	0.00		5.40
	(-1.01)				(-2.25)	(**3.16**)		
4	-0.04						0.02	0.46
	(-0.91)						(0.51)	
Panel B: q = 12								
1	-0.20	2.18						16.29
	(-0.56)	(**2.65**)						
2	-0.47	4.22	-0.02	-0.06				34.76
	(-0.99)	(**3.76**)	(**-3.81**)	(-0.25)				
3	-0.28				-0.00	-0.00		1.23
	(-0.68)				(-0.02)	(-0.22)		
4	-0.34						0.14	1.64
	(-0.74)						(0.27)	
Panel C: q = 60								
1	1.89	-3.78						23.31
	(2.21)	(-0.74)						
2	-2.02	-7.56	-0.03	-2.56				79.34
	(-2.28)	(-2.00)	(**-4.62**)	(**-4.45**)				
3	1.07				0.14	-0.20		47.35
	(*1.65*)				(**4.88**)	(*-1.73*)		
4	1.90						0.27	21.81
	(*1.72*)						(0.30)	
Panel D: Overall assessment								
1	+	+						
2	-	+/-	-	-				
3	+				+/-	+/-		
4	+						(+)	

Table 16 shows the respective results with empirical factors as predictors. The variable SMB^* predicts a decreasing market volatility at $q = 1$ (significant within FF3 and PS) and at $q = 12$ (significant within all models). The slope of HML^* for $q = 1$ is significantly positive within

C, while it is significantly negative within FF5. For $q = 12$, HML^* forecasts $SVAR$ significantly negatively within FF5 and for $q = 60$ its slope is significantly negative within both FF3 and FF5. Moreover, $CUMD$ is a significant predictor for increasing volatility at the one-month horizon. On the other hand, its 60-month forecast is significantly negative. Furthermore, the slope of CL is significantly negative for $q = 1$ and $q = 12$. Finally, the slope of $TERM$ is significantly positive for the 60-month horizon, while DEF significantly predicts increasing volatility at the one- and 12-month horizons.

The overall assessment with regard to all seven forecasting horizons is presented in Panel D of Table 16. Overall, SMB^* predicts market volatility significantly negatively within FF3 and FF5. In contrast, the signs of the slopes of SMB^* within C and PS are indeterminate: Within C, the slope is significantly negative for the 12-month horizon, whereas it is significantly positive for the 36-month horizon (available upon request); within PS, the predictive slope is significantly negative for the one-, three-, and 12-month horizons, whereas it is significantly positive for the 36-month horizon (available upon request). The slope of HML^* is assessed as significantly negative within FF3, PS, and FF5. Thereby, the evaluation for PS is due to the fact that the slope of HML^* within PS is significantly negative for the 24- and 36-month horizons, which are not tabulated here. On the contrary, we assess the sign of the slope of HML^* within C as indeterminate. The reason for this assessment is that the slope is significantly positive for the one- and three-month horizons, while it is significantly negative for the 36- and 48-month horizons. The same applies to the slope of $CUMD$: It is significantly positive for the one- and three-month horizons but significantly negative for the 48- and 60-month horizons and is therefore assessed as indeterminate. Finally, CL is assessed as a significant predictor of decreasing volatility within PS, while $TERM$ and DEF forecast volatility significantly positively within FF5. Altogether, these results are in line with the US evidence.

Taken as a whole, all candidate ICAPM state variables as well as all candidate state variables constructed from empirical factors forecast at least one moment of future aggregate returns within each combination of variables on the right-hand side (though not at each forecasting horizon). Consequently, the results suggest that each of the multifactor models investigated satisfies ICAPM criterion 2, that is, the candidate state variables are related to changes in the investment opportunity set and predict at least one moment of the distribution of aggregate returns.

Table 16. Multiple predictive regressions for state variables constructed from empirical factors (*SVAR*).

This table shows the results for our multiple predictive regressions with the variance ($SVAR$) of the European stock market as the left-hand side variable for horizons of one month (Panel A), 12 months (Panel B), and 60 months (Panel C). The right-hand side variables are the current European values of the size premium SMB^{*}, the value premium HML^{*}, the momentum premium $CUMD$, the liquidity factor CL, the slope of the yield curve $TERM$, and the corporate bond default spread DEF. The original sample period is November 1990 to December 2011. However, one, 12, and 60 observations are lost in each of the one-, 12-, and 60-month horizon regressions and the original sample period for the regressions containing $CUMD$ or CL is November 1995 to December 2011. The first row of each regression shows the estimated slope coefficients and the second row displays Newey–West t-ratios calculated with one, 12, or 60 lags (in parentheses). The levels of statistical significance of the estimated slopes are indicated by italics (10% level), an underline (5% level), and boldface (1% level). The R^2 (%) value presents the *adjusted* R^2 (%) value. Panel D shows the overall assessment of the variables' predictive characteristics, jointly considering the regression results for forecasting horizons of one month, three months, 12 months, 24 months, 36 months, 48 months, and 60 months. A + sign (- sign) means that the variable all in all forecasts positive (negative) market returns, while +/- indicates that the sign of the respective predictive slope is assessed as indeterminate. Parentheses indicate that the forecast is insignificant for all forecasting horizons.

Row	SMB*	HML*	CUMD	CL	TERM	DEF	R² (%)
Panel A: q = 1							
1	-0.00	0.00					2.05
	(**-2.75**)	(1.06)					
2	-0.00	0.00	0.01				6.28
	(-1.61)	(**2.63**)	(<u>2.19</u>)				
3	-0.00	0.00		-0.07			6.47
	(<u>-2.56</u>)	(1.38)		(**-2.79**)			
4	-0.00	-0.00			-0.01	0.37	19.42
	(-0.92)	(<u>-2.15</u>)			(-0.72)	(**3.53**)	
Panel B: q = 12							
1	-0.02	-0.00					9.95
	(<u>-2.26</u>)	(-0.54)					
2	-0.02	0.00	-0.01				4.28
	(*-1.83*)	(0.54)	(-0.37)				
3	-0.02	-0.00		-1.00			23.74
	(<u>-1.98</u>)	(-0.28)		(<u>-2.14</u>)			
4	-0.01	-0.01			-0.20	2.51	28.57
	(*-1.65*)	(**-2.79**)			(-0.62)	(**3.33**)	
Panel C: q = 60							
1	0.01	-0.02					17.20
	(0.33)	(<u>-2.20</u>)					
2	0.03	-0.01	-0.09				14.16
	(1.20)	(-1.59)	(**-2.84**)				
3	0.03	-0.00		-0.20			8.07
	(0.94)	(-0.88)		(-0.15)			
4	-0.01	-0.02			1.76	-3.16	35.90
	(-0.16)	(<u>-2.17</u>)			(**2.94**)	(-0.61)	
Panel D: Overall assessment							
1	-	-					
2	+/-	+/-	+/-				
3	+/-	-		-			
4	-	-			+	+	

5.2.2. Generalized method of moments results

I first regard the generalized method of moments (GMM) estimation results of the ICAPM specifications, which are reported in Table 17. Panel A shows that, on the one hand, the CAPM, HL, and KLVN fail to price 25 stock portfolios sorted by size and book-to-market (SBM25), since the MAE values are pretty high and the R^2_{OLS} values are almost zero or even negative. On the other hand, P and CV do a passable job of explaining the average portfolio returns and CV performs best (with a R^2_{OLS} value of 70% and a MAE value of 0.12%). Regarding Panel B, it seems that the CAPM, HL, and KLVN are also unable to explain the momentum effect represented by 25 stock portfolios sorted by size and momentum (SM25). The results suggest that P explains the momentum effect best (with an R^2_{OLS} value of 65% and a MAE value of 0.18%). The inability of HL and KLVN to explain the SBM25 and SM25 mean excess returns is not in line with the US results of MSC, where the models show R^2 values of 74% and 77% (over SBM25), as well as 50% and 62% (over SM25), respectively. It seems that most models have more difficulties explaining the size/momentum portfolios than explaining the size/book-to-market portfolios, which coincides with the US results. Overall, it seems that all models perform (at least slightly) better within the US sample of MSC than within our European data set. Regarding the estimates of the market (covariance) risk price (ICAPM criterion 1), the estimates for HL (insignificant for SBM25 and significant for SM25) and KLVN (for SBM25 and SM25, but the latter without significance) seem to be a reasonable estimate of the relative risk aversion of the representative investor (RRA), with values between one and 10.

To assess whether the multifactor models satisfy ICAPM criteria 3a and 3b, I compare the overall results of the predictive regressions (Panel D of Tables 13 and 15) with the signs of the models' estimated risk prices. First, I examine the results for SBM25. Regarding HL, γ_{TERM} shows a negative sign, which is at odds with the evidence from our predictive regressions that $TERM$ predicts positive market returns within HL. Consequently, HL does not satisfy ICAPM criterion 3a. Moreover, the positive sign of γ_{DEF} does not agree with the result from the predictive regressions that DEF forecasts market volatility positively within HL. As a result, ICAPM criterion 3b is also not met by HL.

In the case of P, γ_{TERM} and γ_{DEF} both show positive signs, whereas γ_{DY} and γ_{RF} are negative. Thus, the signs of γ_{DEF}, γ_{DY}, and γ_{RF} do not agree with the evidence from the predictive regressions that DEF forecasts negative aggregate returns, whereas DY and RF both forecast positive aggregate returns. Hence, ICAPM criterion 3a is not satisfied by P. In addition, the negative signs of γ_{DY} and γ_{RF} are at odds with the negative slopes on DY and RF in the time-series regressions with aggregate volatility on the left-hand side. Consequently, ICAPM criterion 3b is not satisfied by P.

Regarding CV, γ_{TERM}, γ_{PE}, and γ_{VS} all show a negative sign. Thus CV does not meet ICAPM criterion 3a, since the signs of both γ_{TERM} and γ_{VS} contradict the positive slope of $TERM$ and the positive slope of VS in the return-predicting regressions. However, ICAPM criterion 3b is satisfied by CV if one interprets the indeterminate signs of the slopes associated with PE and VS in the volatility-predicting regressions as positive.

Finally, I examine KLVN, where both γ_{TERM} and γ_{CP} show negative signs. The sign of γ_{TERM} conflicts with the evidence from the predictive regressions that $TERM$ forecasts positive returns. Consequently, my results suggest that ICAPM criterion 3a is not met by KLVN. However, KLVN satisfies criterion 3b, since the negative signs of γ_{TERM} and γ_{CP} agree with the positive slope of $TERM$ and the positive slope of CP in the volatility-predicting regressions.

I conduct the same analysis for SM25 and conclude that HL satisfies ICAPM criterion 3a if the indeterminate sign of the slope of DEF within the return-predicting regressions associated with HL is interpreted as positive. Moreover, P satisfies both ICAPM criteria 3a and 3b if the indeterminate sign of the slope of DEF within the volatility-predicting regressions associated with P is interpreted as positive. Furthermore, CV meets criterion 3b if the indeterminate sign of the slope of PE (VS) within the volatility-predicting regressions associated with CV is taken as positive (negative). However, KLVN satisfies neither ICAPM criterion 3a nor criterion 3b, since the negative (positive) sign of γ_{TERM} (γ_{CP}) does not agree with the positive (negative) sign of the slope of $TERM$ (CP) in the return-predicting regressions or with the positive sign of the slope of CP in the volatility-predicting regressions. Yet only HL and KLVN show an RRA estimate that is between one and 10, so only HL can be justified by the ICAPM.

All in all, when I regard investment opportunities that are driven by changing expected aggregate returns, of all the models with economic factors, only HL can be justified by the ICAPM, since the model meets ICAPM criteria 1, 2, and 3a. However, this result holds only for SM25. When I consider investment opportunities that are driven by changing market volatility, KLVN seems to be justifiable as an application of the ICAPM within my European data set, since the model meets ICAPM criteria 1, 2, and 3b (although only when tested over SBM25). Nonetheless, the explanatory power of both models is very weak within my sample. In contrast, the US results of MSC indicate that none of the models with economic factors can be justified by the ICAPM.

Table 18 shows the results for the multifactor models that include empirical factors. Here C seems to have the highest explanatory power for both SBM25 and SM25, while FF3, PS, and FF5 do a reasonably good job of explaining the SBM25 mean excess returns. However, all four models seem to have more difficulties explaining SM25 than explaining SBM25 and FF3 and PS seem to completely fail in explaining the momentum effect. Those results coincide with the US evidence of MSC, although, overall, the models' explanatory powers seem to be slightly better according to the US results than according to my European evidence. Regarding the estimates of the market (covariance) risk price, it seems that all four models meet ICAPM criterion 1 when the models are tested over SBM25 (although the RRA estimate associated with FF3 is insignificant). Looking at the SM25 results, it seems that FF3, C, and PS meet criterion 1.

Table 17. Factor risk premiums for ICAPM specifications.
This table reports the estimation results for the multifactor models with economically motivated factors using first-stage GMM. Panel A displays the results for the 25 size/book-to-market portfolios of the European stock market as test assets and Panel B presents the results for the 25 size/momentum portfolios. I add the market excess return to both sets of test assets. The market (covariance) risk price is represented by γ and γ_{TERM}, γ_{DEF}, γ_{DY}, γ_{RF}, γ_{PE}, γ_{VS}, and γ_{CP} denote the (covariance) risk prices associated with the slope of the yield curve, the corporate bond default spread, the market dividend-to-price ratio, the short-term risk-free rate, the aggregate price–earnings ratio, the value spread, and the Cochrane–Piazzesi factor, respectively. The estimates of the (covariance) risk prices are reported in the first row that corresponds to each model and the second row contains the asymptotic GMM robust t-statistics (in parentheses). The levels of statistical significance of the estimated (covariance) risk prices are indicated by italics (10% level), an underline (5% level), and boldface (1% level). The second to last column shows the average absolute pricing errors (%) and the last column presents the ordinary least squares (OLS) cross-sectional R^2 value of each model. The sample period is December 1990 to December 2011.

Model	γ	γ_{TERM}	γ_{DEF}	γ_{DY}	γ_{RF}	γ_{PE}	γ_{VS}	γ_{CP}	MAE (%)	R^2_{OLS}
Panel A: SBM25										
CAPM	2.33								0.24	-0.07
	(1.50)									
HL	3.21	-63.54	50.89						0.23	-0.02
	(1.63)	(-0.54)	(0.38)							
P	-14.69	132.94	594.22	-6.99	-363.24				0.15	0.50
	(-1.20)	(0.68)	(*1.76*)	(*-1.71*)	(*-1.90*)					
CV	80.87	-141.60				-72.62	-16.22		0.12	0.70
	(<u>1.97</u>)	(-0.85)				(*-1.87*)	(-1.59)			
KLVN	4.08	-18.65						-132.29	0.22	0.03
	(<u>2.23</u>)	(-0.13)						(*-1.79*)		
Panel B: SM25										
CAPM	2.37								0.36	-0.27
	(1.52)									
HL	6.78	285.68	505.58						0.33	-0.08
	(<u>2.46</u>)	(*1.85*)	(1.53)							
P	29.80	45.23	-147.01	8.87	347.50				0.18	0.65
	(<u>2.42</u>)	(0.34)	(-0.67)	(<u>2.16</u>)	(*1.75*)					
CV	27.05	-62.42				-23.05	13.51		0.30	0.15
	(1.12)	(-0.47)				(-1.04)	(<u>1.96</u>)			
KLVN	1.21	-31.54						100.50	0.36	-0.25
	(0.59)	(-0.28)						(0.99)		

Next, I want to assess whether the models with empirical risk factors satisfy ICAPM criteria 3a and 3b. First, I consider the SBM25 results. In the case of FF3, γ_{SMB} is negative, whereas γ_{HML} is significant and positive. The negative sign of γ_{SMB} is at odds with the positive sign of the slope of SMB^* in the return-predicting regressions. Thus, FF3 does not satisfy criterion 3a when tested over SBM25. Moreover, the negative sign of γ_{SMB} disagrees with my evidence that SMB^* forecasts aggregate volatility negatively. Consequently, ICAPM criterion 3b is not met by FF3 either.

Regarding C, γ_{SMB} is negative, whereas $\gamma_{HML}$and γ_{UMD} are both significantly positive. These results coincide with the evidence that SMB^* forecasts negative market returns and HML^* as well as UMD^* both forecast positive market returns. Consequently, ICAPM criterion 3a is satisfied by C. In addition, criterion 3b is met by C if one interprets the overall indeterminate signs of the slopes of SMB^*, HML^*, and UMD^* within the volatility-predicting regressions as positive, negative, and negative, respectively.

Regarding PS, γ_{SMB} is positive, whereas γ_L is negative, at odds with my evidence that SMB^* forecasts negative market returns and CL predicts positive market returns. Moreover, the sign of γ_L disagrees with the fact that CL forecasts aggregate volatility negatively. Thus, neither ICAPM criterion 3a nor criterion 3b is satisfied by PS.

In the case of FF5, γ_{SMB} and γ_{DEF} are both positive, which does not agree with the negative slopes of SMB^* and DEF within the return-predicting regressions. Moreover, the positive signs of γ_{TERM} and γ_{DEF} disagree with the evidence that both $TERM$ and DEF predict volatility positively. Thus, neither ICAPM criterion 3a nor 3b is met by FF5 when the model is tested over SBM25.

I repeat the analysis for SM25 and conclude that neither model satisfies ICAPM criterion 3a. To be specific, the negative slope of γ_{HML} within FF3 does not coincide with the positive slope of HML^* in the return-predicting regressions. Moreover, the positive slope of γ_{SMB} within C does not agree with the negative slope of SMB^* within the return-predicting regressions. Furthermore, the signs of all risk prices of PS and FF5 do not agree with the signs of the slopes within the respective return-predicting regressions.

With regard to ICAPM criterion 3b, I see that the risk prices of FF3, PS, and FF5 do not satisfy all the sign restrictions implied by the volatility-predicting regressions associated with these three models. In particular, the negative sign of γ_{HML} does not coincide with the negative sign of the slope of HML^* within FF3. Moreover, the negative signs of γ_{HML} and γ_L do not agree with the negative signs of the slopes of HML^* and CL, respectively, within PS, while the negative and positive signs of γ_{HML} and γ_{DEF} do not agree with the negative and positive signs of the slopes of HML^* and DEF, respectively, within FF5. However, C satisfies ICAPM criterion 3b if the overall indeterminate signs of the slopes of SMB^*, HML^*, and UMD^* within the volatility-predicting regressions associated with C are all assumed to be negative.

All in all, my results indicate that, of all the multifactor models with empirical risk factors, only C can be justified by the ICAPM within my European sample. Since the ICAPM criteria 1, 2, 3a (only when tested over SBM25), and 3b (for both sets of test assets) are all satisfied by C, this result holds for investment opportunities driven by changing expect market returns, as well as by changes in the second moment of aggregate returns. This result agrees with the US results of MSC to some extent, where C satisfies the first ICAPM criterion, as well as criteria 2 and 3a over both sets of test assets. However, MSC show that FF3 satisfies all the ICAPM criteria within their US sample when tested over SBM25, which contradicts my European results.

Table 18. Factor risk premiums for empirical risk factors.
This table reports the estimation results for the multifactor models with empirical factors using first-stage GMM. Panel A displays the results for the 25 size/book-to-market portfolios of the European stock market as test assets and Panel B presents the results for the 25 size/momentum portfolios. I add the market excess return to both sets of test assets. The market (covariance) risk price is represented by γ and γ_{SMB}, γ_{HML}, γ_{UMD}, γ_{L}, γ_{TERM}, and γ_{DEF} denote the (covariance) risk prices associated with the size factor, the value factor, the momentum factor, the liquidity factor, the slope of the yield curve, and the corporate bond default spread, respectively. The estimates of the (covariance) risk prices are reported in the first row that corresponds to each model and the second row contains the asymptotic GMM robust t-statistics (in parentheses). The levels of statistical significance of the estimated (covariance) risk prices are indicated by italics (10% level), an underline (5% level), and boldface (1% level). The second to last column shows the average absolute pricing errors (%) and the last column presents the OLS cross-sectional R^2 value of each model. The sample period is December 1990 to December 2011.

Model	γ	γ_{SMB}	γ_{HML}	γ_{UMD}	γ_{L}	γ_{TERM}	γ_{DEF}	MAE (%)	R^2_{OLS}
Panel A: SBM25									
FF3	1.89	-2.96	7.32					0.13	0.50
	(1.21)	(-1.00)	(**3.11**)						
C	7.22	-1.37	11.25	14.00				0.10	0.74
	(**3.01**)	(-0.40)	(**3.66**)	(**2.95**)					
PS	7.64	1.57	12.86		-556.52			0.12	0.61
	(<u>2.22</u>)	(0.31)	(<u>2.31</u>)		(*-1.74*)				
FF5	6.80	3.74	9.51			64.49	456.41	0.14	0.60
	(<u>2.33</u>)	(0.75)	(**2.91**)			(0.49)	(<u>2.06</u>)		
Panel B: SM25									
FF3	3.30	2.26	-13.57					0.32	-0.08
	(*1.81*)	(0.66)	(*-1.82*)						
C	4.47	0.27	8.13	6.97				0.16	0.64
	(<u>2.51</u>)	(0.09)	(*1.88*)	(**3.45**)					
PS	7.85	8.56	-12.19		-428.95			0.33	-0.03
	(**2.64**)	(<u>2.06</u>)	(-1.57)		(<u>-2.52</u>)				
FF5	20.17	46.85	-8.68			-842.76	1096.86	0.25	0.39
	(*1.70*)	(<u>2.00</u>)	(-0.58)			(<u>-2.57</u>)	(1.57)		

Overall, the multifactor models with empirical risk factors seem to show, on average, greater explanatory power than the models including economically motivated state variables. Nevertheless, CV and P keep up with the empirically motivated models, especially P, which seems to perform as well as C in explaining the size/momentum portfolios. Table 19 augments Table 3 with my results regarding each model's consistency with the ICAPM.

Table 19. Consistency of multifactor models with the ICAPM.
This table shows whether the ICAPM criteria of MSC are satisfied for the European stock market by each multifactor model investigated. The criteria are described in Sec. 2.2.1.7. A check mark (✓) means that the respective criterion is satisfied. A checkmark with an asterisk (✓*) for criterion 3a or 3b means that the respective criterion is satisfied if one assumes the overall sign of the slope of one or more state variables in the respective predictive regressions that are actually indeterminate to have the correct sign. ICAPM criterion 2 is satisfied by all models and is independent of the testing assets. Therefore, it is not tabulated here. The corresponding US results of MSC are additionally shown in parentheses (MSC, Table 1). My sample period is December 1990 to December 2011, whereas the sample period of MSC is July 1963 to December 2008. Panel A displays the results for the 25 size/book-to-market portfolios of the European stock market as test assets and Panel B presents the results for the 25 size/momentum portfolios.

	ICAPM criterion 1	ICAPM criterion 3a	ICAPM criterion 3b
Panel A: SBM25			
HL	✓ (x)	x (x)	x (x)
P	x (x)	x (x)	x (x)
CV	x (x)	x (x)	✓* (x)
KLVN	✓ (x)	x (✓)	✓ (x)
FF3	✓ (✓)	x (✓)	x (✓)
C	✓ (✓)	✓ (✓)	✓* (x)
PS	✓ (x)	x (✓)	x (x)
FF5	✓ (x)	x (x)	x (✓)
Panel B: SM25			
HL	✓ (✓)	✓* (x)	x (x)
P	x (x)	✓ (x)	✓* (x)
CV	x (x)	x (x)	✓* (x)
KLVN	✓ (x)	x (x)	x (x)
FF3	✓ (x)	x (x)	x (x)
C	✓ (✓)	x (✓)	✓* (x)
PS	✓ (x)	x (x)	x (x)
FF5	x (x)	x (x)	x (x)

5.2.3. Results of the sensitivity analysis

Appendix B displays the results of the sensitivity analysis described in Sec. 4.2.3. In the first step, I analyze whether these methodological adjustments result in noticeable changes with regard to the models' abilities to explain the cross section of average excess returns. First, the results in Appendix B show that an intercept added to the models' pricing equations is economically and statistically significant within all models and over both the SBM25 and SM25 portfolios (varying between 1% and 2% per month). The sole exception is the intercept added to C in the test over SM25. The addition of an intercept leads to all models showing higher R^2_{OLS} and lower MAE values than in the original tests. Especially with regard to SM25, I observe that the R^2_{OLS} value of the CAPM, HL, CV, KLVN, FF3, and PS increase substantially. These findings indicate that these models are not specified correctly. Regarding SBM25, the highest R^2_{OLS} value is now shown for CV (0.81, followed by C with 0.77), while the highest R^2_{OLS} value over SM25 is still obtained with P (0.67, followed by C with 0.65).

Second, a re-estimation of the models with second-stage GMM hardly affects the models' explanatory powers, now measured by the weighted least squares coefficient of determination. Third, the addition of excess bond returns to the test assets does not change the measures of most models' explanatory powers that much, even though the pricing of both stock and bond risk premiums at the same time should be more demanding for the models. The greatest changes are shown by the explanatory ratios of FF3 over SBM25 ($R^2_{OLS} = 0.29$) and by C over both SBM25 ($R^2_{OLS} = 0.48$) and SM25 ($R^2_{OLS} = 0.49$). Hence, both models seem to have noticeably more difficulties in simultaneously explaining stock and bond risk premiums. Consequently, the highest R^2_{OLS} over SBM25 is now shown for CV (0.74) and the highest R^2_{OLS} over SM25 is shown for P (0.65). Fourth, the R^2_{OLS} and MAE estimates from the original tests also seem to be quite robust with respect to excluding the market excess return from the test assets. Hence, the impact of forcing the models to price one of the factors seems to be rather low.

Fifth, while the estimation of the asset pricing models in expected return–beta form via OLS has no impact on the original R^2_{OLS} and MAE estimates, most models' abilities to explain the cross section of returns change noticeably when I use generalized least squares (GLS). With respect to the ICAPM specifications, the explanatory powers (now measured by the GLS coefficient of determination, R^2_{GLS}) of HL and CV especially (tested over SBM25 and SM25), as well as of P (estimated over SM25), improve considerably in comparison to the benchmark first-stage GMM tests. Regarding the models with empirical risk factors, I note that the explanatory measures associated with FF3, PS, and FF5 decrease over SBM25 but increase for SM25, while the explanatory ratio of C increases over SBM25 but decreases over SM25 (compared to the R^2_{OLS} value obtained with first-stage GMM). Overall, the highest explanatory power for SBM25 is now shown for CV ($R^2_{GLS} = 0.91$), while both P and CV now have the greatest ability to price SM25 ($R^2_{GLS} = 0.99$). In comparison, the R^2_{GLS} value of C is 0.88 over SBM25 and just 0.49 over SM25. Finally, the use of different proxies for the state variable innovations, that is, innovations obtained from AR(1) and VAR(1) processes, instead of first differences has a rather small impact on the models' abilities to price the test portfolios. If I employ a VAR(1) process, it is remarkable that the R^2_{OLS} value of P increases to 0.62 while that of CV decreases to 0.55 (regarding SBM25) and the R^2_{OLS} of P decreases to 0.51 with regard to SM25.

In the second step, I evaluate whether the methodological adjustments result in noticeable changes with respect to the fulfillment of the ICAPM criteria described in Sec. 2.2.1.7. Regardless of how I change the predictive regressions, ICAPM criterion 2 is still met by each model, since the candidate state variables significantly forecast either expected market returns or market volatility in at least one of the predictive regressions associated with each model. The results with regard to the satisfaction of the remaining ICAPM criteria are summarized in Table 20.

Table 20. Sensitivity analysis.
This table shows how the consistency of each multifactor model with MSC's ICAPM criteria changes when one modifies the underlying empirical methodology. The criteria are described in Sec. 2.2.1.7. The first column displays the methodological change. The first to third columns corresponding to each model shows the consistency with ICAPM criteria 1, 3a, and 3b, respectively. A check mark (✓) means that the respective criterion is satisfied. A checkmark and an asterisk (✓*) for criterion 3a or criterion 3b means that the respective criterion is met if one assumes the overall sign of the slope of one or more state variables in the respective predictive regressions that are actually indeterminate to have the correct sign. The grayed out areas indicate that the respective criterion is generally independent of the respective methodological adjustment. ICAPM criterion 2 is satisfied by all the models and changes in methodology. Therefore, it is not tabulated here. Panel A displays the results for the 25 size/book-to-market portfolios of the European stock market as test assets and Panel B presents the results for the 25 size/momentum portfolios.

Model	HL			P			CV			KLVN			FF3			C			PS			FF5		
ICAPM criterion	1	3a	3b	1	3a	3b	1	3a	3b	1	3a	3b	1	3a	3b	1	3a	3b	1	3a	3b	1	3a	3b
Panel A: SBM25																								
Original assessment	✓	x	x	x	x	x	x	x	✓*	✓	x	✓	✓	x	x	✓	✓	✓*	✓	x	x	✓	x	x
Employing alternative proxies for the state variables associated with SMB and HML														✓	✓		✓*	✓*		x	x		✓*	✓*
Using alternative volatility measures (Beeler and Campbell, 2012)			✓*			x			✓*			✓*			x			x			x			x
Using alternative volatility measures (Bansal et al., 2005)			✓*			x			✓*			x			x			x			x			x
Using alternative volatility measures (squared return)			x			x			x			x			x			x			x			x
Including an intercept in the pricing equations	x	x	x	x	x	x	x	x	✓*	x	x	x	x	x	x	✓	✓	✓*	x	x	x	x	x	x
Estimating the asset pricing models with second-stage GMM	✓	✓*	x	✓	x	x	x	x	x	✓	✓	x	✓	x	x	x	✓	✓*	x	x	x	x	x	x
Adding excess bond returns to the test assets	✓	x	✓	x	x	x	x	x	✓*	✓	x	✓	✓	x	x	✓	✓	✓*	x	x	x	✓	x	x
Excluding the market excess return from the test assets	✓	x	✓	x	x	x	x	x	✓*	✓	✓	x	✓	x	x	✓	✓	✓*	✓	x	x	✓	x	x
Estimating the asset pricing models in expected return–beta form (OLS)	✓	x	x	✓	x	x	✓	x	✓*	✓	x	✓	✓	x	x	✓	✓	✓*	✓	x	x	✓	x	x
Estimating the asset pricing models in expected return–beta form (GLS)	✓	✓*	x	✓	x	x	✓	x	✓*	✓	x	x	✓	x	x	✓	✓	✓*	✓	x	x	✓	x	x
Estimating the asset pricing models with orthogonal factors	✓	x	x	✓	x	x	✓	x	✓*	✓	x	✓	✓	x	x	✓	✓	✓*	✓	x	x	✓	x	x
Using different proxies for the state variable innovations (AR(1))	✓	x	✓	x	x	x	x	x	✓*	✓	✓	x												
Using different proxies for the state variable innovations (VAR(1))	✓	x	x	x	x	x	x	x	✓*	✓	x	✓												

Table 20 (continued)

Model	HL			P			CV			KLVN			FF3			C			PS			FF5		
ICAPM criterion	1	3a	3b	1	3a	3b	1	3a	3b	1	3a	3b	1	3a	3b	1	3a	3b	1	3a	3b	1	3a	3b
Panel B: SM25																								
Original assessment	✓	✓*	x	x	✓	✓*	x	x	✓*	✓	x	x	✓	x	x	✓	x	✓*	✓	x	x	x	x	x
Employing alternative proxies for the state variables associated with SMB and HML														x	x		x	x		x	x		x	x
Using alternative volatility measures (Beeler and Campbell, 2012)			x			✓*			✓*			✓*			x			✓*			x			x
Using alternative volatility measures (Bansal et al., 2005)			x			x			✓*			✓			x			✓			x			x
Using alternative volatility measures (squared return)			x			x			x			✓			✓			x			✓*			✓*
Including an intercept in the pricing equations	x	✓*	x	x	x	✓*	x	x	✓*	x	x	x	x	x	x	✓	✓	✓*	x	x	x	✓	x	x
Estimating the asset pricing models with second-stage GMM	x	✓*	x	x	✓	✓*	x	x	x	✓	x	x	✓	x	x	✓	✓	✓*	✓	x	x	x	x	x
Adding excess bond returns to the test assets	✓	x	x	x	x	x	x	x	✓*	✓	x	✓	✓	x	x	✓	✓	✓*	x	x	x	x	x	x
Excluding the market excess return from the test assets	✓	✓*	x	x	x	x	x	x	✓*	x	x	x	✓	x	x	✓	x	✓*	✓	x	x	x	x	x
Estimating the asset pricing models in expected return–beta form (OLS)	✓	✓*	x	✓	x	x	✓	x	✓*	✓	x	x	✓	x	x	✓	✓	✓*	✓	x	x	✓	x	x
Estimating the asset pricing models in expected return–beta form (GLS)	✓	✓*	x	✓	x	x	✓	x	x	✓	✓	x	✓	x	x	✓	x	✓*	✓	x	x	✓	x	x
Estimating the asset pricing models with orthogonal factors	✓	✓*	x	✓	✓	✓*	✓	x	✓*	✓	x	x	✓	x	x	✓	x	✓*	✓	x	x	✓	x	x
Using different proxies for the state variable innovations (AR(1))	✓	✓*	x	x	✓	✓*	x	x	✓*	x	x	x												
Using different proxies for the state variable innovations (VAR(1))	✓	✓*	x	x	x	x	x	x	✓*	✓	x	x												

I consider the outcomes for the SBM25 portfolios, which are displayed in Panel A of Table 20. First, I observe that the overall results for P, FF3, C, PS, and FF5 are hardly affected by the methodological changes. In particular, although P shows a market (covariance) risk price that is economically reasonable as an estimate of the RRA in some of the conducted robustness checks (e.g., when the model is estimated in expected return–beta form), the signs of the estimated risk prices remain inconsistent with the signs of the slopes from the predictive regressions (criteria 3a and 3b) in all tests. Consequently, my results strongly indicate that P is not justifiable as an application of the ICAPM within the European stock market when tested over SBM25. The results for FF3 are quite similar, but with one exception: When I use proxies for the state variables associated with SMB and HML that are constructed in the same way as CL and $CUMD$, respectively, FF3 meets all three ICAPM criteria and is thus justifiable by the ICAPM. In all other robustness checks, neither criterion 3a nor criterion 3b is satisfied by FF3. Thus, FF3's consistency with the ICAPM seems to rely on how the state variables associated with SMB and HML are constructed. In all but one of the robustness checks, C remains consistent with the ICAPM. The sole change in methodology that results in an inconsistency of C with the ICAPM is the use of second-stage GMM, where the RRA estimate is slightly above 10 (10.55). However, note that the model's consistency with criterion 3b is quite vulnerable: If I employ alternative volatility measures, the criterion is no longer met and in all the remaining robustness checks the fulfillment of the criterion depends on how one interprets the signs of several predictive regression slopes that are actually indeterminate (indicated by ✓* in the Table 20). Nevertheless, these results suggest that C can be interpreted as an application of the ICAPM within the European stock market when tested over SBM25. Like the P model, PS does not meet criterion 3a or 3b in any of the robustness checks. Hence, the results strongly indicate that PS is also not justifiable by the ICAPM when the model is estimated over the European SBM25 portfolios. The outcomes for FF5 are quite similar to the FF3 results: The model meets both ICAPM criteria 3a and 3b when I employ alternative proxies for the state variables associated with SMB and HML. However, the signs of the slopes of these two alternative proxies are both indeterminate in the return-predicting regressions associated with FF5, which also applies to the slopes of the alternative proxy for the state variable associated with SMB and the slope of DEF within the respective volatility-predicting regressions. All other methodological changes result in an inconsistency of FF5 with the ICAPM.

Second, note that the assessment of the success or failure of HL, CV, and KLVN to meet the ICAPM criteria is more sensitive with respect to the underlying empirical methodology. To be specific, HL satisfies criterion 3a in two of the robustness checks and criterion 3b in five of them. Consequently, HL becomes consistent with the ICAPM in seven tests (although, in some instances, the assessment depends on how one interprets the indeterminate signs of several predictive regression slopes) but remains inconsistent with the ICAPM in five robustness checks, so one can hardly draw the overall conclusion that the model is justifiable by Merton's theory. CV becomes consistent in three of the robustness checks, that is, when I estimate the model in expected return–beta form or with orthogonal factors, since the market (covariance) risk price becomes economically reasonable as an estimate of the RRA in these tests, ranging from 3.76 to 5.66. However, criterion 3a is not satisfied by CV in any of the tests. Finally, KLVN remains justifiable by the ICAPM in eight of the 13 robustness checks, but whether the model meets criterion 3a or 3b changes quite often.

I now look at the results associated with the SM25 portfolios. Note that the overall outcomes are quite stable for most models. In particular, HL remains consistent with the ICAPM in all but three robustness checks. Thereby, ICAPM criterion 3b remains unmet in all tests. When an intercept is included in the model's pricing equation or when the model is estimated with second-stage GMM, the market (covariance) price of risk associated with HL becomes economically implausible as an estimate of the RRA (the RRA estimates are -1.92 and 10.65). Moreover, criterion 3a is no longer met when excess bond returns are added to the set of test assets. The results associated with P are somewhat more sensitive with regard to the underlying empirical methodology. Whether or not the three ICAPM criteria are met by the model changes quite often from test to test. Nevertheless, the overall conclusion is stable: P remains inconsistent with the ICAPM in all but one robustness check. The model only becomes justifiable by the ICAPM if I orthogonalize the risk factors with respect to the market excess return, which is done within Petkova's (2006) original paper on the model.

In two of the robustness checks, CV satisfies the ICAPM restrictions, that is, when I estimate the model in expected return–beta form via OLS and when I estimate it with orthogonalized factors. In all other tests, either criterion 1 remains unmet by the model or criterion 3b is no longer satisfied while the model remains inconsistent with criterion 3a for all tests. KLVN remains inconsistent with the ICAPM criteria 3a and 3b, unless I employ alternative volatility measures in the volatility-predicting regressions or add excess bond returns to the testing portfolios (which both lead to criterion 3b being fulfilled by the model) or when I estimate the model in expected return–beta representation using GLS (such that criterion 3a is met). In all other robustness checks, neither criterion 3a nor criterion 3b is satisfied by KLVN.

Observe that FF3 becomes consistent with the ICAPM in only one of the robustness checks, that is, when the volatility-predicting regressions are conducted using the monthly squared continuously compounded market return as the left-hand side variable, so that the model fulfills criterion 3b. None of the other tests result in FF3 meeting criterion 3a or 3b. Regarding C, one can see that the model remains consistent with the ICAPM in all but two robustness checks. The sole two tests where the model meets neither criterion 3a nor criterion 3b are when I construct the state variables associated with SMB and HML in the same way as CL and $CUMD$, respectively, and when I employ squared returns in the volatility-predicting regressions.

The results for PS are very stable. The original assessment regarding the fulfillment of the ICAPM criteria changes in only three robustness checks: When I use the squared market return as the volatility measure, the model satisfies criterion 3b, and when I add excess bond returns to SM25 or include an intercept in the model's pricing equation, the RRA estimates become economically implausible (10.98 and -7.28). Finally, observe that FF5 remains unjustifiable by the ICAPM in all robustness checks.

5.3. Results for the Cost of Equity Capital for European Industries

In the following I present my results with respect to (multi-)factor asset pricing models and their ability to estimate the cost of equity capital (CE) for European industries (Research Questions 4 and 5).[38]

5.3.1. Risk loadings

5.3.1.1. Risk loadings from full-period regressions

Table 21 displays the results for the estimation of regression (4.14) over the full sample period, for each industry as well as the average estimates across all 35 industries. The adjusted R^2 values averaged across industries seem to be comparable to the values obtained by Fama and French (1997, FF97) but higher than those found by Gregory and Michou (2009, GM09). The lowest value is obtained using the CAPM (60%) and the highest is obtained using FF3, C, PS, and FF5 (65%). Regarding the adjusted R^2 values of individual industries, the lowest is obtained by estimating the CAPM over *Tobacco* (14%), whereas the highest value is obtained for C and *Banks* (88%). Overall, it seems that industries with a low adjusted R^2 within one model tend to have a low adjusted R^2 within the other models, confirming the findings of GM09.

For the CAPM, FF3, and C, whose factors are all excess returns, the regression intercepts represent the pricing errors (alphas) and thus provide a hint of whether the industries' expected returns are appropriately described by these asset pricing models. The alphas averaged across industries are slightly above zero for all three models (13, 11, and 15 basis points per month), which indicates that the expected excess return on the average (typical) industry seems to be well (but not perfectly) described by the three models. The number of statistically significant alphas across industries is lowest for the CAPM (two positive ones), whereas FF3 shows eight significant alphas and C four. The industry with the highest CAPM alpha is *Tobacco* (1.11% per month), while the industry with the highest FF3 alpha as well as the highest C alpha is *Software & Computer Services* (1.09% and 1.17% per month, respectively). Moreover, the loading on SMB averaged across industries is between 0.12 and 0.13, that on HML is between 0.10 and 0.12, and that on UMD is -0.03, while the typical industry shows, unsurprisingly, a market beta close to one.

[38] This chapter is based on Sec. 4 of Lutzenberger (2014c).

Table 21. Full-period regressions.

This table shows for each model the intercept (*a*) (in percent), the intercept's t-statistic, the slope coefficients, and the adjusted R^2 value from an OLS time-series regression over the full sample period August 1990 to December 2011. The variables *b*, *s*, *h*, *w*, and *l* represent the slope coefficients of the market excess return, size factor, value factor, momentum factor, and liquidity factor, respectively; *m*, *d*, *y*, *r*, *p*, *v*, and *c* represent the slope coefficients on the innovations in the slope of the yield curve, corporate bond default spread, market dividend-to-price ratio, short-term risk-free rate, aggregate price–earnings ratio, value spread, and Cochrane–Piazzesi factor, respectively. The fourth to last row shows the averages of the regression coefficients and of the adjusted R^2 values across all industries. The third to last row displays the regression coefficients' average standard errors across all industries. The standard errors are heteroskedasticity consistent, following White (1980). Finally, Sig +ve (Sig -ve) represents the number of industries for which the respective regression coefficient is significantly positive (negative) at the 5% significance level. The industries are defined in Table 9.

	CAPM				FF3						C						
	a (%)	*t(a)*	*b*	Adj. R^2	*a* (%)	*t(a)*	*b*	*s*	*h*	Adj. R^2	*a* (%)	*t(a)*	*b*	*s*	*h*	*w*	Adj. R^2
Ind1	0.37	1.57	0.84	0.52	0.28	1.15	0.85	0.13	0.23	0.53	0.29	1.01	0.85	0.13	0.22	-0.01	0.53
Ind2	0.30	1.63	0.99	0.73	0.13	0.70	0.98	-0.07	0.22	0.74	0.10	0.54	0.98	-0.07	0.23	0.02	0.74
Ind3	-0.20	-0.71	0.99	0.52	-0.46	-1.63	1.00	0.25	0.58	0.58	-0.30	-1.05	0.97	0.24	0.54	-0.14	0.58
Ind4	0.00	-0.01	1.36	0.61	0.08	0.25	1.45	0.94	0.47	0.69	0.10	0.32	1.45	0.94	0.47	-0.02	0.69
Ind5	0.57	1.42	1.13	0.41	0.60	1.40	1.17	0.40	0.21	0.43	0.58	1.25	1.17	0.40	0.22	0.01	0.42
Ind6	-0.12	-0.71	1.08	0.78	-0.30	-2.06	1.11	0.37	0.52	0.85	-0.18	-1.26	1.09	0.37	0.49	-0.10	0.86
Ind7	0.15	0.57	1.01	0.56	0.08	0.29	1.03	0.23	0.25	0.57	0.20	0.70	1.00	0.23	0.22	-0.10	0.57
Ind8	0.00	0.02	1.07	0.77	-0.12	-0.69	1.08	0.15	0.28	0.79	-0.20	-1.14	1.09	0.15	0.30	0.07	0.79
Ind9	0.08	0.32	1.38	0.74	0.42	1.68	1.42	0.27	-0.36	0.77	0.60	2.18	1.38	0.27	-0.40	-0.15	0.77
Ind10	0.13	0.71	1.15	0.77	0.19	1.00	1.21	0.54	0.26	0.81	0.28	1.50	1.19	0.54	0.23	-0.07	0.82
Ind11	-0.05	-0.31	0.98	0.77	-0.07	-0.46	1.02	0.40	0.29	0.82	0.01	0.10	1.00	0.40	0.27	-0.07	0.82
Ind12	0.13	0.83	0.99	0.76	0.31	2.06	1.03	0.40	-0.02	0.79	0.31	2.02	1.03	0.40	-0.03	0.00	0.79
Ind13	-0.08	-0.25	1.15	0.55	-0.19	-0.57	1.16	0.05	0.20	0.55	-0.13	-0.39	1.14	0.05	0.19	-0.05	0.55
Ind14	0.38	1.80	0.63	0.44	0.24	1.13	0.63	0.08	0.26	0.46	0.17	0.76	0.65	0.08	0.28	0.06	0.46
Ind15	0.35	1.89	0.53	0.42	0.18	0.94	0.51	-0.13	0.18	0.44	0.04	0.21	0.54	-0.13	0.21	0.11	0.45
Ind16	0.16	0.79	0.85	0.62	0.16	0.79	0.89	0.33	0.20	0.65	0.19	0.94	0.88	0.33	0.19	-0.03	0.65
Ind17	-0.07	-0.33	1.09	0.70	-0.02	-0.11	1.13	0.39	0.18	0.72	0.10	0.45	1.10	0.39	0.15	-0.10	0.72
Ind18	0.45	1.89	0.83	0.53	0.12	0.55	0.77	-0.47	0.21	0.58	0.00	0.01	0.80	-0.46	0.24	0.10	0.58
Ind19	1.11	3.48	0.43	0.14	0.84	2.55	0.39	-0.26	0.25	0.16	0.67	1.85	0.43	-0.26	0.29	0.14	0.16
Ind20	0.30	1.73	0.72	0.60	0.42	2.49	0.75	0.22	-0.06	0.61	0.34	1.91	0.76	0.22	-0.04	0.07	0.61
Ind21	0.51	2.35	0.48	0.30	0.37	1.79	0.42	-0.47	-0.09	0.36	0.25	1.15	0.45	-0.47	-0.06	0.10	0.37
Ind22	0.12	0.62	0.68	0.52	0.08	0.37	0.67	-0.08	0.01	0.51	0.13	0.60	0.66	-0.08	0.00	-0.04	0.51
Ind23	0.17	0.88	0.85	0.62	0.14	0.66	0.85	0.02	0.06	0.62	0.30	1.36	0.82	0.02	0.03	-0.13	0.63
Ind24	-0.07	-0.33	1.09	0.67	0.41	1.78	1.14	0.33	-0.55	0.74	0.34	1.62	1.15	0.33	-0.53	0.06	0.74
Ind25	-0.24	-1.20	1.05	0.70	-0.03	-0.15	1.12	0.63	0.06	0.76	-0.02	-0.11	1.12	0.63	0.06	-0.01	0.75
Ind26	0.16	0.60	0.96	0.52	0.46	1.91	0.93	-0.40	-0.71	0.63	0.32	1.38	0.96	-0.40	-0.68	0.11	0.63
Ind27	0.29	1.73	0.66	0.58	0.19	1.15	0.67	0.15	0.24	0.60	0.08	0.42	0.70	0.15	0.27	0.10	0.61
Ind28	0.20	1.11	0.71	0.57	-0.13	-0.77	0.66	-0.32	0.31	0.64	-0.17	-0.93	0.67	-0.32	0.32	0.04	0.64
Ind29	-0.19	-1.10	1.24	0.81	-0.60	-4.18	1.22	-0.14	0.55	0.86	-0.33	-2.00	1.16	-0.14	0.49	-0.23	0.88
Ind30	-0.19	-0.85	1.16	0.71	-0.60	-2.97	1.10	-0.44	0.37	0.76	-0.31	-1.43	1.04	-0.45	0.30	-0.24	0.78
Ind31	-0.25	-0.98	1.30	0.69	-0.63	-2.48	1.27	-0.16	0.48	0.72	-0.12	-0.48	1.16	-0.18	0.36	-0.42	0.77
Ind32	-0.06	-0.30	0.72	0.53	-0.21	-1.26	0.76	0.49	0.55	0.68	-0.17	-0.91	0.75	0.49	0.54	-0.04	0.68
Ind33	-0.04	-0.27	0.94	0.82	-0.02	-0.17	0.98	0.38	0.22	0.86	-0.01	-0.07	0.98	0.38	0.21	-0.01	0.86
Ind34	0.27	0.74	1.36	0.55	1.09	3.37	1.43	0.47	-0.98	0.65	1.17	3.47	1.42	0.47	-1.00	-0.07	0.65
Ind35	-0.02	-0.05	1.63	0.61	0.45	1.21	1.61	-0.33	-0.94	0.67	0.65	1.66	1.57	-0.34	-0.99	-0.17	0.68
Means	0.13		0.97	0.60	0.11		0.98	0.12	0.11	0.65	0.15		0.97	0.12	0.10	-0.03	0.65
Mean SE	0.00		0.06		0.00		0.06	0.11	0.10		0.00		0.06	0.11	0.09	0.07	
Sig +ve	2		35		4		35	16	20		3		35	16	21	0	
Sig –ve	0		0		4		0	6	5		1		0	6	5	5	

Table 21 (continued)

	PS							FF5								HL					
	a (%)	*t(a)*	*b*	*s*	*h*	*l*	Adj. R^2	*a* (%)	*t(a)*	*b*	*s*	*h*	*m*	*d*	Adj. R^2	*a* (%)	*t(a)*	*b*	*m*	*d*	Adj. R^2
Ind1	0.27	1.12	0.85	0.14	0.24	-0.77	0.53	0.26	1.10	0.85	0.11	0.23	1.82	1.06	0.54	0.37	1.59	0.83	1.99	-0.32	0.53
Ind2	0.13	0.69	0.98	-0.07	0.22	0.06	0.74	0.12	0.69	0.97	-0.10	0.22	0.59	-0.96	0.74	0.30	1.67	0.98	0.42	-1.09	0.72
Ind3	-0.46	-1.61	1.00	0.25	0.57	0.19	0.58	-0.51	-1.89	1.07	0.32	0.63	2.79	7.35	0.60	-0.22	-0.80	1.01	3.33	3.40	0.54
Ind4	0.08	0.25	1.45	0.94	0.47	-0.05	0.69	0.06	0.20	1.39	0.77	0.42	3.73	-4.60	0.71	0.05	0.17	1.26	5.03	-10.52	0.66
Ind5	0.59	1.37	1.18	0.41	0.22	-0.91	0.42	0.58	1.39	1.11	0.23	0.17	3.50	-4.37	0.44	0.60	1.54	1.07	3.90	-6.30	0.44
Ind6	-0.30	-2.07	1.11	0.37	0.52	-0.26	0.85	-0.31	-2.20	1.14	0.42	0.54	0.45	3.29	0.86	-0.11	-0.69	1.07	1.16	-1.00	0.78
Ind7	0.10	0.37	1.01	0.21	0.22	2.04	0.57	0.08	0.27	1.07	0.30	0.28	-0.27	3.44	0.57	0.15	0.55	1.01	0.24	0.75	0.55
Ind8	-0.12	-0.69	1.08	0.15	0.28	0.03	0.79	-0.11	-0.68	1.07	0.13	0.27	-0.03	-0.98	0.79	0.02	0.10	1.04	0.19	-2.66	0.78
Ind9	0.43	1.73	1.41	0.26	-0.38	1.28	0.77	0.40	1.63	1.41	0.22	-0.37	1.69	-0.53	0.77	0.08	0.32	1.37	2.07	-0.60	0.75
Ind10	0.19	1.03	1.20	0.53	0.25	0.58	0.81	0.18	0.97	1.20	0.50	0.25	0.99	-0.75	0.81	0.16	0.86	1.11	1.83	-4.51	0.78
Ind11	-0.08	-0.52	1.02	0.41	0.30	-0.76	0.82	-0.06	-0.39	1.04	0.48	0.31	-1.87	1.60	0.82	-0.04	-0.23	0.96	-1.05	-2.24	0.77
Ind12	0.31	2.01	1.04	0.40	-0.02	-0.38	0.79	0.31	2.02	1.05	0.41	-0.02	0.20	1.21	0.79	0.14	0.86	0.98	0.90	-1.18	0.76
Ind13	-0.20	-0.60	1.17	0.06	0.22	-1.08	0.55	-0.23	-0.74	1.27	0.25	0.29	0.90	11.36	0.58	-0.13	-0.43	1.23	1.32	8.94	0.57
Ind14	0.22	1.04	0.65	0.10	0.28	-1.75	0.46	0.25	1.17	0.61	0.05	0.24	-0.01	-1.99	0.46	0.39	1.89	0.60	0.07	-3.07	0.44
Ind15	0.15	0.83	0.53	-0.11	0.21	-2.23	0.45	0.19	1.01	0.55	-0.02	0.21	-2.51	2.62	0.46	0.34	1.87	0.55	-2.54	2.04	0.45
Ind16	0.15	0.74	0.90	0.34	0.21	-0.96	0.65	0.17	0.81	0.91	0.40	0.22	-1.32	1.74	0.66	0.16	0.83	0.84	-0.65	-1.34	0.62
Ind17	-0.02	-0.11	1.13	0.39	0.18	-0.06	0.72	-0.03	-0.13	1.08	0.29	0.15	1.47	-3.28	0.72	-0.04	-0.20	1.03	1.96	-5.48	0.71
Ind18	0.11	0.50	0.78	-0.46	0.23	-1.12	0.58	0.14	0.61	0.77	-0.46	0.21	-1.24	-1.20	0.58	0.44	1.88	0.84	-2.01	0.81	0.54
Ind19	0.80	2.41	0.43	-0.23	0.30	-3.85	0.17	0.87	2.62	0.40	-0.21	0.25	-2.86	-0.59	0.17	1.11	3.52	0.43	-3.21	-0.22	0.15
Ind20	0.42	2.47	0.75	0.22	-0.06	-0.07	0.61	0.42	2.48	0.76	0.25	-0.05	-0.08	1.62	0.61	0.30	1.71	0.72	0.34	0.31	0.60
Ind21	0.37	1.78	0.42	-0.47	-0.09	0.05	0.36	0.38	1.81	0.46	-0.38	-0.06	-1.50	2.76	0.37	0.48	2.29	0.52	-2.15	5.22	0.34
Ind22	0.08	0.38	0.67	-0.08	0.01	0.31	0.51	0.08	0.38	0.68	-0.04	0.03	-0.76	1.15	0.51	0.11	0.59	0.69	-0.82	1.27	0.52
Ind23	0.14	0.63	0.85	0.02	0.07	-0.54	0.62	0.15	0.69	0.84	0.02	0.06	-0.46	-0.74	0.61	0.18	0.91	0.84	-0.43	-1.03	0.62
Ind24	0.43	1.91	1.12	0.31	-0.58	2.40	0.74	0.40	1.75	1.14	0.31	-0.55	0.62	0.18	0.74	-0.08	-0.34	1.09	1.15	0.17	0.67
Ind25	-0.02	-0.09	1.11	0.62	0.05	1.21	0.76	-0.03	-0.17	1.16	0.72	0.09	-0.99	3.73	0.76	-0.24	-1.18	1.04	0.24	-0.84	0.70
Ind26	0.47	1.92	0.92	-0.41	-0.72	1.12	0.63	0.46	1.92	0.88	-0.49	-0.74	0.66	-4.14	0.63	0.16	0.58	0.97	-0.17	1.23	0.52
Ind27	0.19	1.14	0.68	0.15	0.25	-0.24	0.60	0.21	1.26	0.69	0.21	0.25	-2.00	0.51	0.61	0.30	1.81	0.65	-1.65	-1.56	0.58
Ind28	-0.14	-0.82	0.67	-0.32	0.32	-0.76	0.64	-0.12	-0.70	0.69	-0.24	0.33	-2.07	1.70	0.65	0.19	1.08	0.73	-2.48	2.00	0.59
Ind29	-0.61	-4.24	1.22	-0.13	0.56	-0.72	0.86	-0.59	-4.16	1.20	-0.14	0.54	-0.90	-1.56	0.86	-0.18	-1.03	1.22	-1.15	-2.50	0.81
Ind30	-0.61	-2.98	1.10	-0.44	0.37	-0.46	0.76	-0.60	-3.08	1.15	-0.35	0.40	-0.95	3.93	0.76	-0.21	-0.99	1.20	-1.53	4.64	0.72
Ind31	-0.61	-2.42	1.26	-0.17	0.46	1.45	0.72	-0.60	-2.44	1.22	-0.23	0.44	-1.25	-5.38	0.73	-0.22	-0.89	1.25	-1.65	-5.47	0.70
Ind32	-0.22	-1.35	0.77	0.50	0.56	-1.30	0.68	-0.19	-1.15	0.76	0.53	0.54	-2.38	-1.37	0.70	-0.03	-0.13	0.66	-1.49	-6.27	0.56
Ind33	-0.03	-0.25	0.99	0.39	0.23	-0.92	0.86	-0.01	-0.12	1.01	0.45	0.23	-1.27	1.88	0.86	-0.03	-0.21	0.93	-0.51	-1.53	0.82
Ind34	1.09	3.40	1.43	0.47	-0.98	0.15	0.65	1.07	3.35	1.40	0.36	-1.01	2.96	-2.16	0.66	0.27	0.76	1.35	3.57	-0.94	0.55
Ind35	0.47	1.28	1.59	-0.35	-0.97	2.31	0.67	0.42	1.12	1.61	-0.39	-0.94	3.14	1.14	0.68	-0.06	-0.15	1.68	2.49	6.54	0.61
Means	0.11		0.98	0.13	0.11	-0.17	0.65	0.11		0.99	0.13	0.12	0.02	0.50	0.65	0.14		0.97	0.25	-0.67	0.61
Mean SE	0.00		0.06	0.11	0.10	1.25		0.00		0.06	0.11	0.10	0.95	2.05		0.00		0.06	0.97	2.10	
Sig +ve	4		35	15	22	0		4		35	16	20	5	4		2		35	7	1	
Sig -ve	4		0	6	5	1		4		0	6	5	6	1		0		0	6	5	

Table 21 (continued)

	P								CV							KLVN					
	a (%)	*t(a)*	*b*	*m*	*d*	*y*	*r*	Adj. R^2	*a* (%)	*t(a)*	*b*	*m*	*p*	*v*	Adj. R^2	*a* (%)	*t(a)*	*b*	*m*	*c*	Adj. R^2
Ind1	0.31	1.29	1.06	2.64	-1.69	0.08	1.36	0.53	0.33	1.21	0.89	2.02	-0.06	-0.04	0.53	0.37	1.58	0.84	2.24	-0.31	0.53
Ind2	0.42	2.25	0.82	1.33	0.20	-0.05	2.03	0.73	0.27	1.18	1.02	0.49	-0.03	-0.01	0.72	0.30	1.64	0.99	0.35	0.19	0.72
Ind3	-0.05	-0.18	0.83	4.79	4.99	-0.06	3.23	0.54	-0.67	-1.81	1.67	3.12	-0.67	-0.16	0.56	-0.20	-0.71	0.98	2.69	0.58	0.53
Ind4	0.12	0.38	1.20	5.75	-9.99	-0.02	1.58	0.66	-0.22	-0.53	1.65	5.73	-0.30	0.01	0.63	-0.01	-0.03	1.35	6.21	-0.68	0.64
Ind5	0.67	1.59	0.98	4.36	-5.58	-0.03	1.03	0.43	0.07	0.12	1.84	4.33	-0.70	-0.01	0.43	0.57	1.42	1.13	4.53	-0.30	0.43
Ind6	-0.02	-0.13	0.79	0.58	0.74	-0.09	-1.19	0.79	-0.49	-1.91	1.62	1.24	-0.53	-0.10	0.80	-0.12	-0.72	1.08	1.32	-0.12	0.78
Ind7	0.12	0.43	1.09	0.32	0.29	0.02	0.17	0.55	0.36	0.71	0.71	0.18	0.30	-0.04	0.55	0.15	0.57	1.01	0.18	0.01	0.55
Ind8	0.09	0.52	0.85	-0.05	-1.40	-0.07	-0.47	0.78	-0.15	-0.60	1.29	0.37	-0.22	0.01	0.77	0.00	-0.01	1.07	1.07	-0.98	0.78
Ind9	0.23	0.98	1.23	3.68	0.72	-0.05	3.55	0.75	0.13	0.35	1.29	2.11	0.08	0.11	0.75	0.08	0.34	1.37	1.23	1.21	0.75
Ind10	0.27	1.46	0.96	2.58	-3.32	-0.05	1.67	0.78	-0.04	-0.14	1.40	2.13	-0.24	-0.05	0.77	0.13	0.72	1.15	1.90	0.31	0.77
Ind11	-0.01	-0.04	0.76	-2.04	-1.19	-0.07	-2.10	0.78	0.03	0.15	0.87	-0.91	0.11	-0.06	0.77	-0.05	-0.30	0.98	-1.03	0.18	0.77
Ind12	0.07	0.40	1.09	0.45	-2.02	0.04	-1.02	0.76	0.57	1.86	0.37	0.96	0.61	-0.01	0.76	0.13	0.83	0.99	0.95	0.04	0.76
Ind13	0.00	0.00	1.16	3.02	9.84	-0.02	3.72	0.57	-0.36	-0.91	1.55	0.75	-0.39	0.03	0.55	-0.08	-0.25	1.15	0.44	0.41	0.55
Ind14	0.45	2.14	0.55	0.66	-2.60	-0.02	1.31	0.44	0.41	0.96	0.59	0.26	0.04	-0.08	0.44	0.38	1.81	0.63	0.19	0.11	0.43
Ind15	0.30	1.57	0.61	-2.79	1.60	0.02	-0.55	0.44	0.29	1.14	0.62	-2.68	-0.08	-0.09	0.46	0.35	1.94	0.53	-2.66	-0.03	0.44
Ind16	0.12	0.61	0.93	-0.74	-1.92	0.03	-0.23	0.62	0.15	0.54	0.87	-0.56	-0.01	-0.06	0.63	0.16	0.79	0.86	-0.62	0.08	0.62
Ind17	0.03	0.15	0.95	2.56	-4.77	-0.03	1.32	0.71	0.10	0.29	0.84	2.32	0.23	-0.02	0.70	-0.07	-0.36	1.09	2.90	-0.80	0.70
Ind18	0.52	2.08	0.65	-2.11	2.02	-0.06	-0.17	0.54	0.84	2.14	0.28	-2.08	0.55	0.01	0.54	0.44	1.91	0.84	-1.37	-0.96	0.54
Ind19	0.93	2.82	0.65	-4.65	-2.02	0.07	-3.18	0.16	1.44	3.28	-0.02	-3.21	0.45	-0.14	0.17	1.11	3.52	0.44	-2.64	-0.76	0.15
Ind20	0.29	1.64	0.77	0.54	0.04	0.02	0.43	0.60	0.27	1.07	0.76	0.32	-0.04	0.00	0.60	0.29	1.71	0.72	0.70	-0.53	0.60
Ind21	0.43	2.08	0.56	-2.72	4.85	0.01	-1.24	0.33	1.03	3.10	-0.26	-2.52	0.73	-0.02	0.32	0.50	2.38	0.49	-1.63	-1.20	0.32
Ind22	0.07	0.37	0.79	-0.79	0.60	0.03	0.04	0.52	0.48	1.72	0.17	-0.92	0.51	-0.01	0.52	0.12	0.63	0.68	-0.80	-0.14	0.52
Ind23	0.18	0.83	0.68	-1.67	-0.29	-0.05	-2.65	0.63	0.13	0.45	0.92	-0.36	-0.07	-0.05	0.62	0.17	0.89	0.85	-0.47	0.15	0.62
Ind24	-0.19	-0.86	1.28	0.60	-1.21	0.06	-1.24	0.67	-0.50	-1.91	1.69	1.16	-0.59	0.06	0.68	-0.07	-0.33	1.09	1.02	0.17	0.67
Ind25	-0.31	-1.41	1.10	-0.44	-1.41	0.02	-1.49	0.70	-0.47	-1.57	1.37	0.30	-0.31	0.01	0.70	-0.24	-1.19	1.04	-0.09	0.52	0.70
Ind26	-0.14	-0.50	1.43	-1.85	-2.22	0.15	-3.76	0.55	-0.08	-0.17	1.30	-0.24	-0.33	0.08	0.52	0.17	0.61	0.96	-0.34	0.12	0.52
Ind27	0.31	1.90	0.56	-2.18	-1.09	-0.03	-1.13	0.58	0.47	2.06	0.41	-1.55	0.25	-0.06	0.59	0.29	1.75	0.67	-1.40	-0.20	0.58
Ind28	0.26	1.41	0.59	-2.31	2.95	-0.05	0.39	0.60	0.05	0.19	0.93	-2.60	-0.21	-0.03	0.59	0.20	1.15	0.71	-2.77	0.22	0.59
Ind29	0.06	0.28	0.74	-0.73	0.78	-0.16	1.03	0.83	0.06	0.26	0.89	-1.00	0.35	-0.08	0.81	-0.19	-1.10	1.24	-1.31	0.45	0.81
Ind30	-0.02	-0.08	0.75	-1.61	7.63	-0.15	-0.06	0.75	0.04	0.15	0.83	-1.85	0.32	-0.01	0.72	-0.19	-0.85	1.16	-2.03	0.26	0.72
Ind31	0.05	0.19	0.66	-1.44	-1.47	-0.20	0.61	0.73	-0.02	-0.06	0.98	-1.30	0.32	-0.06	0.69	-0.25	-0.97	1.29	-1.95	0.91	0.69
Ind32	0.03	0.17	0.43	-2.30	-4.90	-0.08	-1.70	0.58	0.08	0.29	0.54	-1.09	0.18	-0.14	0.56	-0.06	-0.32	0.72	-0.70	-0.52	0.54
Ind33	-0.06	-0.42	0.92	-1.10	-1.62	0.00	-1.28	0.82	-0.04	-0.18	0.95	-0.42	-0.01	-0.07	0.82	-0.04	-0.28	0.95	-0.16	-0.35	0.82
Ind34	0.14	0.37	1.75	4.45	-3.45	0.14	1.81	0.56	0.02	0.04	1.68	3.65	-0.32	0.21	0.57	0.28	0.79	1.34	2.12	2.10	0.56
Ind35	-0.10	-0.24	1.92	3.62	5.23	0.08	2.41	0.62	0.30	0.44	1.15	2.05	0.46	0.15	0.61	-0.03	-0.07	1.63	2.89	-1.14	0.61
Means	0.16		0.92	0.30	-0.33	-0.02	0.12	0.62	0.14		0.96	0.29	0.01	-0.02	0.61	0.13		0.97	0.31	-0.03	0.61
Mean SE	0.00		0.14	1.09	2.24	0.04	1.61		0.00		0.42	0.98	0.41	0.05		0.00		0.06	1.11	0.60	
Sig +ve	5		35	9	3	2	0		4		23	7	0	0		2		35	4	2	
Sig -ve	0		0	8	3	7	2		0		0	5	0	7		0		0	3	2	

The standard error of the market beta averaged across industries is 0.06 for all models except P and CV (for which it is 0.14 and 0.42, respectively) and those of SMB, HML, and UMD are 0.11, 0.09 to 0.10, and 0.07, respectively. These values appear to be relatively low, although they are slightly higher than those obtained by FF97. These findings suggest that the slopes are precisely estimated to some degree. In contrast, the standard errors of the loadings on L, $\Delta TERM$, ΔDEF, ΔRF, ΔPE, and ΔCP averaged across industries are considerably higher. Those of the slopes of ΔDY and ΔVS are exceptions, but the magnitudes of the slope estimates on these two factors are themselves considerably lower. For instance, the standard error of the loading on L averaged across industries is 1.25 and that of the slope of ΔDEF even ranges between 2.05 and 2.24. These values indicate that the loadings on these risk factors are not estimated precisely. Thus, the results so far suggest that only the CAPM, FF3, and C risk loadings can be estimated somewhat accurately, whereas one is unable to obtain reliable estimates of the remaining multifactor models' slopes.

5.3.1.2. Implied standard deviations of true risk loadings

Table 22 presents the results of the implied standard deviations of the true factor risk loadings, which are estimated following Eq. (4.15). The estimated implied standard deviations indicate that the true factor risk loadings of the industries vary considerably through time, which confirms the findings of FF97 and GM09. For instance, the lower and upper bounds of the typical industry's current true CAPM beta, which I calculate as the average 60-month rolling market slope across time and all industries minus/plus twice the slope's implied standard deviation averaged across industries, are 0.65 and 1.29, respectively. The current true CAPM beta of the typical industry might be anywhere within this range. With regard to the FF3 and C results, one can see that the average volatility of the true market slope declines slightly (from 0.16 within the CAPM to 0.13 and 0.12 within FF3 and C, respectively), in line with the FF97 study. The lower and upper bounds of the typical industry's true SMB and HML loadings are (-0.03; 0.35) and (-0.42; 0.60) within FF3, respectively. The corresponding limits of the UMD slope within C are (-0.28; 0.18).

Regarding the results for individual industries, it seems that some industries behave like small stocks at some times but like large stocks at other times, as noted by FF97. For instance, the lower and upper bounds of *Forestry & Paper*'s true loading on SMB within FF3 are (-0.42; 1.04). When the industry's true loading on SMB is near one, the industry behaves like a typical small stock, while the industry behaves like a typical large stock in periods when the loading is negative (Fama and French, 1996). Similarly, some industries seem to behave like value stocks in some periods but like growth stocks in other periods. For example, the lower and upper bounds of the true loading of *Oil & Gas* on HML within FF3 are (-1.19; 1.35). The industry behaves like a typical growth stock when the current true loading is clearly negative and a typical value stock when the true loading is within the top range of this interval (Fama and French, 1996).

Table 22. Implied standard deviations of true factor risk loadings in 60-month rolling regressions.

This table displays the implied standard deviations of the true factor risk loadings estimated with 60-month rolling OLS time-series regressions. The variables *b*, *s*, *h*, *w*, and *l* represent the slope coefficients of the market excess return, size factor, value factor, momentum factor, and liquidity factor, respectively; *m*, *d*, *y*, *r*, *p*, *v*, and *c* represent the slope coefficients of the innovations in the slope of the yield curve, corporate bond default spread, market dividend-to-price ratio, short-term risk-free rate, aggregate price–earnings ratio, value spread, and Cochrane–Piazzesi factor, respectively. The fourth to last row displays the average implied standard deviations across all industries. Moreover, lower bound (upper bound) represents the average slope coefficient across all months and industries minus (plus) twice the slope coefficient's average implied standard deviation across all industries. Finally, the last row (%tge=0) displays the percentage of industries that show an implied standard deviation equal to zero, that is, for which the average estimation error variance exceeds the time-series variance. The industries are defined in Table 9. The original sample period is August 1990 to December 2011 and 59 months are lost when the implied standard deviations are computed.

	CAPM	FF3			C				PS				FF5					HL			P					CV				KLVN		
	b	*b*	*s*	*h*	*b*	*s*	*h*	*w*	*b*	*s*	*h*	*l*	*b*	*s*	*h*	*m*	*d*	*b*	*m*	*d*	*b*	*m*	*d*	*y*	*r*	*b*	*m*	*p*	*v*	*b*	*m*	*c*
Ind1	0.13	0.12	0.00	0.63	0.14	0.01	0.46	0.41	0.07	0.00	0.62	2.40	0.12	0.10	0.62	0.00	0.00	0.12	0.00	0.00	0.00	1.88	0.00	0.00	0.00	0.00	0.00	0.00	0.11	0.12	0.00	0.00
Ind2	0.00	0.00	0.00	0.22	0.05	0.00	0.22	0.15	0.00	0.00	0.22	1.61	0.00	0.00	0.23	0.00	0.00	0.00	0.00	0.00	0.02	0.50	0.00	0.00	3.36	0.00	0.00	0.00	0.00	0.00	0.00	0.22
Ind3	0.00	0.00	0.37	0.00	0.03	0.36	0.00	0.00	0.00	0.36	0.00	6.09	0.00	0.15	0.00	2.20	0.00	0.11	3.10	0.00	0.00	3.04	0.00	0.00	0.00	0.00	3.29	0.00	0.06	0.00	3.19	0.00
Ind4	0.18	0.00	0.00	0.25	0.05	0.00	0.00	0.14	0.00	0.00	0.25	0.00	0.00	0.00	0.35	0.00	0.00	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.59	0.00	0.87	0.16	0.19	0.00	0.00
Ind5	0.12	0.12	0.00	0.52	0.00	0.00	0.38	0.41	0.11	0.00	0.55	0.00	0.17	0.00	0.61	0.00	1.34	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.15	0.18	0.00	2.56
Ind6	0.18	0.05	0.00	0.00	0.10	0.04	0.00	0.17	0.06	0.05	0.00	3.52	0.08	0.00	0.00	0.65	4.78	0.21	0.74	4.38	0.18	1.02	4.91	0.00	0.00	0.62	0.76	0.50	0.04	0.19	1.04	0.00
Ind7	0.10	0.16	0.00	0.38	0.12	0.00	0.44	0.00	0.17	0.03	0.42	6.44	0.16	0.17	0.36	0.00	3.94	0.11	0.00	4.68	0.20	0.00	4.99	0.08	0.00	0.00	0.00	0.00	0.05	0.09	0.00	0.00
Ind8	0.15	0.07	0.15	0.00	0.06	0.16	0.06	0.08	0.05	0.14	0.00	1.98	0.05	0.12	0.10	1.01	0.00	0.12	0.98	1.46	0.17	0.24	2.46	0.00	0.00	0.46	0.39	0.62	0.00	0.15	0.70	0.00
Ind9	0.31	0.32	0.34	0.00	0.28	0.34	0.00	0.00	0.32	0.34	0.00	0.00	0.32	0.40	0.00	0.65	0.00	0.29	0.00	0.00	0.00	0.27	1.85	0.06	0.00	0.00	0.00	0.32	0.00	0.27	0.00	1.17
Ind10	0.12	0.00	0.00	0.00	0.00	0.00	0.00	0.14	0.00	0.00	0.03	3.89	0.00	0.00	0.00	0.97	2.21	0.10	0.00	1.62	0.31	1.79	2.72	0.07	2.48	0.00	0.87	0.00	0.03	0.13	0.00	0.97
Ind11	0.11	0.02	0.19	0.23	0.04	0.19	0.23	0.08	0.00	0.19	0.23	0.00	0.00	0.15	0.23	0.00	0.00	0.10	0.00	0.00	0.22	1.37	0.00	0.04	1.11	0.00	0.00	0.00	0.04	0.12	0.00	1.46
Ind12	0.10	0.13	0.07	0.13	0.12	0.09	0.09	0.18	0.14	0.04	0.15	3.37	0.12	0.11	0.15	0.00	0.00	0.10	0.00	0.00	0.10	0.00	0.00	0.07	1.42	0.37	0.00	0.33	0.00	0.09	0.00	0.00
Ind13	0.00	0.09	0.34	0.29	0.08	0.36	0.00	0.11	0.03	0.31	0.32	0.00	0.09	0.19	0.23	0.00	2.26	0.00	0.00	3.39	0.00	0.00	5.89	0.00	3.52	0.02	0.80	0.39	0.00	0.00	1.39	0.00
Ind14	0.22	0.19	0.00	0.40	0.18	0.00	0.39	0.12	0.17	0.00	0.38	1.92	0.17	0.00	0.43	0.00	0.00	0.20	0.00	0.00	0.21	1.06	0.00	0.00	1.84	0.76	0.00	0.49	0.04	0.22	1.07	0.64
Ind15	0.19	0.14	0.00	0.44	0.11	0.00	0.41	0.11	0.12	0.00	0.46	1.61	0.12	0.00	0.46	0.00	5.77	0.17	0.00	5.71	0.11	0.00	5.62	0.00	0.00	0.00	0.00	0.00	0.04	0.17	0.00	0.00
Ind16	0.13	0.13	0.00	0.34	0.17	0.00	0.31	0.16	0.13	0.00	0.35	0.00	0.11	0.00	0.40	0.06	8.71	0.14	0.79	8.20	0.17	2.68	8.30	0.00	2.90	0.00	1.42	0.00	0.09	0.14	1.85	0.00
Ind17	0.00	0.00	0.26	0.20	0.01	0.27	0.13	0.08	0.00	0.27	0.25	0.00	0.00	0.23	0.25	1.30	2.62	0.00	1.78	2.68	0.00	1.75	3.32	0.00	2.13	0.69	2.13	0.76	0.11	0.00	2.80	0.00
Ind18	0.21	0.16	0.19	0.20	0.14	0.15	0.25	0.12	0.14	0.21	0.19	0.00	0.16	0.28	0.18	0.00	4.93	0.20	0.00	0.00	0.20	0.00	0.00	0.00	0.00	1.08	0.00	1.11	0.08	0.19	0.00	1.10
Ind19	0.25	0.17	0.12	0.27	0.14	0.05	0.00	0.00	0.16	0.19	0.31	0.00	0.10	0.00	0.36	0.00	0.00	0.20	0.00	0.00	0.33	0.00	0.00	0.00	0.00	0.09	0.00	0.17	0.00	0.20	0.00	0.79

Table 22 (continued)

	CAPM	FF3			C				PS				FF5					HL			P				CV					KLVN		
	b	*b*	*s*	*h*	*b*	*s*	*h*	*w*	*b*	*s*	*h*	*l*	*b*	*s*	*h*	*m*	*d*	*b*	*m*	*d*	*b*	*m*	*d*	*y*	*r*	*b*	*m*	*p*	*v*	*b*	*m*	*c*
Ind20	0.09	0.09	0.00	0.14	0.10	0.00	0.14	0.00	0.10	0.00	0.12	0.00	0.10	0.00	0.15	1.14	0.00	0.10	0.85	0.00	0.13	1.30	0.00	0.00	1.73	0.83	0.62	0.70	0.00	0.09	1.16	0.97
Ind21	0.09	0.09	0.05	0.31	0.03	0.07	0.26	0.14	0.06	0.00	0.28	8.77	0.07	0.11	0.30	0.00	0.00	0.04	0.11	0.00	0.00	1.30	0.00	0.00	0.00	0.00	0.92	0.00	0.02	0.08	0.00	1.27
Ind22	0.00	0.00	0.00	0.46	0.00	0.00	0.47	0.08	0.00	0.00	0.46	3.78	0.00	0.00	0.45	1.06	0.00	0.01	0.44	0.00	0.16	0.00	1.92	0.05	0.00	0.62	1.01	0.58	0.09	0.00	0.63	0.00
Ind23	0.10	0.08	0.08	0.31	0.05	0.10	0.31	0.00	0.09	0.08	0.33	0.66	0.10	0.00	0.31	0.00	3.55	0.09	0.00	4.02	0.29	1.41	4.68	0.08	4.15	0.00	0.00	0.00	0.04	0.11	0.00	0.00
Ind24	0.22	0.16	0.00	0.47	0.21	0.00	0.42	0.19	0.14	0.03	0.48	3.85	0.15	0.00	0.48	0.00	0.00	0.20	0.00	3.56	0.33	0.00	2.88	0.00	0.00	0.73	0.00	0.51	0.09	0.21	0.00	0.00
Ind25	0.04	0.19	0.13	0.22	0.20	0.11	0.20	0.00	0.20	0.15	0.22	0.00	0.16	0.00	0.22	0.00	0.00	0.05	0.00	2.14	0.00	2.41	1.26	0.00	2.72	0.67	0.00	0.55	0.00	0.05	1.49	0.39
Ind26	0.26	0.09	0.26	0.45	0.03	0.24	0.43	0.25	0.09	0.27	0.45	0.00	0.13	0.29	0.41	0.90	2.46	0.26	0.00	0.00	0.41	0.00	0.00	0.00	0.00	0.82	0.00	0.40	0.01	0.26	0.00	0.00
Ind27	0.22	0.23	0.15	0.06	0.24	0.14	0.12	0.00	0.22	0.15	0.07	0.00	0.21	0.00	0.09	0.72	5.24	0.21	1.25	5.49	0.27	1.70	5.42	0.01	0.00	0.48	1.18	0.53	0.00	0.21	1.26	0.00
Ind28	0.13	0.10	0.00	0.04	0.11	0.00	0.06	0.00	0.10	0.00	0.08	0.00	0.09	0.00	0.09	2.21	0.00	0.13	1.56	0.00	0.14	2.10	0.00	0.00	0.00	0.54	1.40	0.42	0.00	0.11	1.21	0.00
Ind29	0.14	0.09	0.00	0.27	0.09	0.06	0.00	0.18	0.10	0.00	0.25	1.02	0.10	0.00	0.25	0.00	0.00	0.15	0.55	0.00	0.08	0.28	0.00	0.00	1.35	0.34	0.93	0.00	0.03	0.14	1.32	0.38
Ind30	0.19	0.34	0.05	0.24	0.29	0.04	0.20	0.18	0.35	0.04	0.23	0.00	0.34	0.04	0.26	1.22	0.00	0.19	0.85	0.00	0.24	0.47	0.00	0.11	0.91	0.00	1.40	0.00	0.03	0.21	2.52	1.51
Ind31	0.31	0.36	0.17	0.46	0.29	0.11	0.32	0.23	0.35	0.20	0.45	0.00	0.35	0.17	0.44	2.30	0.00	0.30	2.03	0.00	0.30	2.51	0.00	0.17	0.00	0.00	2.09	0.00	0.05	0.33	2.89	1.20
Ind32	0.26	0.15	0.00	0.27	0.15	0.00	0.28	0.06	0.15	0.00	0.24	0.00	0.15	0.00	0.21	1.09	2.68	0.23	1.85	0.00	0.21	2.31	0.00	0.04	0.00	0.00	1.64	0.00	0.00	0.26	1.76	0.00
Ind33	0.12	0.06	0.08	0.18	0.07	0.09	0.20	0.00	0.06	0.08	0.18	0.00	0.05	0.04	0.19	0.00	3.34	0.12	0.84	3.00	0.00	1.33	2.98	0.04	2.08	0.00	0.78	0.00	0.00	0.12	1.19	0.00
Ind34	0.50	0.35	0.35	0.14	0.33	0.27	0.12	0.22	0.37	0.36	0.19	0.00	0.35	0.32	0.19	1.25	0.00	0.48	0.00	5.43	0.00	0.00	7.37	0.00	0.00	1.54	0.00	1.67	0.08	0.46	0.00	0.75
Ind35	0.43	0.21	0.00	0.39	0.17	0.00	0.31	0.00	0.20	0.08	0.38	2.53	0.16	0.00	0.42	0.00	0.00	0.40	0.00	0.00	0.50	0.00	0.00	0.00	5.90	0.00	0.00	0.00	0.00	0.42	0.00	0.00
Means	0.16	0.13	0.10	0.25	0.12	0.09	0.21	0.11	0.12	0.10	0.26	1.53	0.12	0.08	0.27	0.54	1.54	0.15	0.51	1.59	0.15	0.93	1.90	0.02	1.07	0.32	0.62	0.31	0.04	0.16	0.78	0.44
Lower bound	0.65	0.76	-0.03	-0.42	0.78	-0.02	-0.35	-0.28	0.78	-0.05	-0.44	-3.34	0.77	-0.01	-0.45	-0.97	-3.01	0.67	-0.76	-3.89	0.64	-1.59	-4.58	-0.06	-1.92	0.30	-0.88	-0.60	-0.09	0.66	-1.34	-0.87
Upper bound	1.29	1.27	0.35	0.60	1.26	0.35	0.47	0.18	1.26	0.36	0.61	2.77	1.26	0.31	0.62	1.17	3.14	1.27	1.27	2.48	1.25	2.15	3.02	0.04	2.37	1.59	1.59	0.65	0.07	1.29	1.80	0.89
%tge=0	14%	17%	49%	14%	9%	40%	23%	31%	20%	40%	11%	54%	20%	54%	11%	54%	60%	9%	57%	60%	31%	37%	54%	66%	57%	49%	51%	49%	37%	14%	51%	57%

Finally, some industries seem to behave like winner stocks at some times but loser stocks at other times. On the other hand, a considerable percentage of industries show zero volatility in their true factor risk loadings. For instance, the implied standard deviation of *Industrial Metals*'s true loading on SMB within FF3 is zero. In addition, the industry's average 60-month rolling slope of SMB within FF3 is 0.85, which indicates that the industry constantly behaves like a typical small stock.

Looking at the results for the remaining multifactor models, it seems that, overall, the percentage of industries that report zero variability of their true factor risk loadings increases to some extent in comparison to the CAPM, FF3, and C. However, the volatilities that are unequal to zero appear to be extremely high for many industries. For instance, although 54% of all industries show zero volatility of their true loading on ΔDEF within P, their mean volatility is 1.9. Thus, the typical industry's current true loading on ΔDEF within P might be anywhere between -4.58 and 3.02. Such a high variation over time is shown by the true loadings on most non-traded factors, with the exceptions of the loadings on ΔDY and ΔVS.

5.3.1.3. Costs of equity forecasts

The results of the CAPM, FF3, and C forecasts of short-term and long-term CEs using full-period and rolling estimates are displayed in Table 23.

Table 23. Summary statistics for CAPM, FF3, and C forecast errors.
This table reports summary statistics for the forecast errors from 60-month rolling and full-period regressions. The first row of each model displays the average mean absolute forecast errors and the second row (in parentheses) reports the average standard deviations of the forecasts errors (in percent). The forecast errors are computed monthly from $t = 60 + q$ to $t = T$, where $t = 60$ corresponds to August 1995, $t = T$ is December 2011, and $q = 1, 12, 24, 36, 48, 60$ months is the forecasting horizon. For the rolling regressions, the forecasting error at t is computed as the realized return at t minus the vector of realized factors at t times the vector of factor slopes that were estimated over $t - q - 59$ to $t - q$. Thus, the intercept from the rolling regressions is not considered. The static (full-period) regressions are estimated from $t = 60 + q$ to $t = T$. The corresponding forecast error at t is computed as the regression's residual at t plus its intercept. For the static regressions, both the average mean absolute forecast errors and the average standard deviations of the forecasts errors are adjusted for degrees of freedom by multiplying by $\frac{T-59-q-1}{T-59-q-K-1}$, where K is the number of factors included in the respective model.

	1 month (%)	1 year (%)	2 years (%)	3 years (%)	4 years (%)	5 years (%)
CAPM (rolling)	2.91	3.03	3.15	3.20	3.20	3.09
	(3.85)	(4.00)	(4.15)	(4.23)	(4.24)	(4.06)
CAPM (static)	2.92	2.96	3.01	3.00	2.94	2.81
	(3.85)	(3.90)	(3.96)	(3.96)	(3.88)	(3.66)
FF3 (rolling)	2.82	3.01	3.16	3.26	3.26	3.14
	(3.69)	(3.97)	(4.20)	(4.35)	(4.37)	(4.15)
FF3 (static)	2.75	2.79	2.83	2.83	2.78	2.68
	(3.60)	(3.64)	(3.69)	(3.69)	(3.60)	(3.43)
C (rolling)	2.82	3.14	3.43	3.52	3.43	3.18
	(3.75)	(4.16)	(4.61)	(4.75)	(4.67)	(4.26)
C (static)	2.74	2.78	2.82	2.81	2.75	2.62
	(3.58)	(3.62)	(3.67)	(3.67)	(3.57)	(3.37)

The magnitudes of the forecast errors and their standard deviations are in line with FF97's findings but smaller than those obtained by GM09. My results indicate that rolling regression loadings are dominated by full-period slopes at all forecasting horizons within all three models, which confirms the findings of FF97 and GM09. In all instances (except the one-month CAPM forecast), the average mean absolute forecast error of the rolling regressions is greater than that of the full-period static regressions. The same applies to the forecast errors' average standard deviations. Moreover, the forecast errors first increase with the horizons, in line with the FF97 results, but then decrease at longer horizons, somewhat in contrast to the FF97 findings.

These results suggest that the manager of a typical industry should prefer full-period risk loadings to rolling regression slopes for both long-term and short-term cash flows. For short-term cash flows, the forecast noise resulting from the increased smoothing of the variation in the true slopes by full-period regressions appears to be offset by their increased precision achieved by the extended estimation period, as noted by FF97 or, for individual firms, Gonedes (1973). My findings concerning longer-horizon cash flows are a sign of the mean reversion of the typical industry's true risk loadings.

Table 24 displays the errors from estimating the observed risk loadings 60 months ahead, using rolling regressions that include past returns, for the CAPM and all eight multifactor models under investigation.

The mean absolute errors of the market and HML slopes seem to be comparable to the GM09 findings for most models, while the mean absolute errors of the SMB and UMD slopes are considerably lower than for GM09. Nevertheless, they show that rolling regressions that use past returns provide rather imprecise estimates for future risk loadings on RM, SMB, HML, and UMD. The errors of the loadings on most of the non-traded factors are markedly worse. For instance, the mean absolute error of the loading on ΔDEF varies between 7.50 and 7.88, whereas the mean absolute error of the loading on RM is only 0.30 within the CAPM. These findings coincide with the regression results from Secs. 5.3.1.1 and 5.3.1.2 and suggest that future risk loadings on non-traded factors cannot be accurately estimated using past returns.

Table 24. Risk loading prediction errors.
This table displays the mean (M), mean absolute (MA), and average standard deviation ($MStd$) of the errors from the prediction of the slope coefficients 60 months ahead, using 60 months of past returns. I calculate the slope prediction error in month t as the slope of a regression involving excess returns and factors from $t + 1$ to $t + 60$ minus the slope of a regression with excess returns and factors from $t - 59$ to t. The variables b, s, h, w, and l represent the slope coefficients of the market excess return, size factor, value factor, momentum factor, and liquidity factor, respectively; m, d, y, r, p, v, and c represent the slope coefficients of the innovations in the slope of the yield curve, corporate bond default spread, market dividend-to-price ratio, short-term risk-free rate, aggregate price–earnings ratio, value spread, and Cochrane–Piazzesi factor, respectively. The original sample period is August 1990 to December 2011 and 119 months are lost when I compute the prediction errors.

	CAPM	FF3			C				PS			
	b	b	s	h	b	s	h	w	b	s	h	l
M	0.01	-0.02	-0.04	0.03	-0.03	-0.04	0.04	0.02	-0.02	-0.04	0.04	0.34
MA	0.30	0.26	0.29	0.49	0.26	0.29	0.48	0.25	0.26	0.30	0.50	4.67
$MStd$	0.33	0.28	0.31	0.52	0.28	0.31	0.52	0.27	0.28	0.32	0.53	5.84

	FF5					HL		
	b	s	h	m	d	b	m	d
M	-0.02	-0.03	0.04	-0.20	0.28	0,00	-0,35	0,66
MA	0.25	0.29	0.50	2.60	7.50	0,29	2,68	7,76
$MStd$	0.27	0.33	0.53	3.01	8.78	0,33	3,15	8,86

	P					CV				KLVN		
	b	m	d	y	r	b	m	p	v	b	m	c
M	-0,01	-0,69	0,52	0,00	-0,48	-0.15	-0.36	0.16	-0.02	0,01	-0,33	-0,06
MA	0,49	3,33	7,88	0,14	4,96	1.28	2.76	1.21	0.11	0,30	3,39	1,86
$MStd$	0,56	3,66	9,00	0,16	5,68	1.41	3.27	1.34	0.13	0,33	3,97	2,08

5.3.2. Risk premiums

Table 25 shows the annualized risk premiums of the traded factors RM, SMB, HML, and UMD, estimated by their sample means (Eq. (4.16)). The annualized estimate of the market risk premium is 4.80%. However, its annualized standard error of 3.55% is fairly high. According to the traditional rule of thumb of plus or minus two standard errors, the true market risk premium might be anywhere between -2.29% and 11.90%. Thus, if one repeated history, one might even obtain a negative average market excess return. Therefore, the premium is not significantly different from zero. The positive historical premium that I obtained might have been the result of chance. For comparison, Fama and French (2012) report a market risk premium of 7.9% p.a., with a standard error of 3.4% for North America (i.e., the US plus Canada) over the sample period of November 1990 to March 2011. The estimated premium for SMB is -5.39%. With a standard error of 1.75% this estimate is statistically significant. Its negative value indicates that large stocks have higher average returns than small stocks in my sample (see Sec. 3.2). The annualized sample mean of HML is 7.76%. Its standard error (1.98%) indicates that it is significantly different from zero. Finally, the estimated risk premium for UMD is 11.21% p.a., which, having a standard error of 3.14%, is significantly positive.

Table 25. Factor risk premiums from sample means.

This table presents the summary statistics for the risk premiums associated with the market excess return, RM; the size factor, SMB; the value factor, HML; and the momentum factor, UMD, obtained from their sample means (in percent). The third and second to last rows present the lower and upper limits, respectively, for the annualized average premiums according to the rule of thumb of plus or minus two standard errors. The last row shows the t-statistic of the mean, that is, the ratio of the mean to its standard error. The sample period is August 1990 to December 2011.

	RM	SMB	HML	UMD
Average premium (monthly, %)	0.40	-0.45	0.65	0.93
Standard error (monthly, %)	0.30	0.15	0.17	0.26
Average premium (annualized, %)	4.80	-5.39	7.76	11.21
Standard error (annualized, %)	3.55	1.75	1.98	3.14
Lower limit (annualized, %)	-2.29	-8.90	3.80	4.94
Upper limit (annualized, %)	11.90	-1.88	11.73	17.48
t-Mean	1.35	-3.07	3.91	3.58

Table 26 displays the annualized results of the risk premium estimation using Fama–MacBeth (1973) regressions with traded factors on the left-hand side (Eqs. (4.17) and (4.18)). At first glance, one can see that the risk premium estimates for a given factor depend on both the asset pricing model in which it is included and the procedure used to estimate the risk loadings, that is, full-period or rolling regressions. Moreover, most factor risk premiums seem to be rather small, ranging somewhere between -1% and 1% p.a. Exceptions are the premiums for ΔDY, ΔPE, ΔVS, and ΔCP. Moreover, I find many relatively high standard errors, rendering most risk premiums not significantly different from zero.

Table 26. Factor risk premiums from cross-sectional regressions with adjusted returns.
This table reports the results from monthly Fama–MacBeth (1973) regressions with the adjusted excess returns on 35 European industries as left-hand side variables. The regressions are conducted under the null hypothesis that the risk premiums for a model's traded factors are equal to their sample means and the industry excess returns are adjusted correspondingly. The right-hand side variables are the factor risk loadings, which are estimated in one static multiple time-series regression (Panel A) or 60-month rolling multiple time-series regressions (Panel B) for each model. The regressions are estimated without an intercept term. The variables λ_L, λ_{TERM}, λ_{DEF}, λ_{DY}, λ_{RF}, λ_{PE}, λ_{VS}, and λ_{CP} represent the estimated risk premiums for the liquidity factor, slope of the yield curve, corporate bond default spread, market dividend-to-price ratio, short-term risk-free rate, aggregate price–earnings ratio, value spread, and Cochrane–Piazzesi factor. All risk premiums are annualized (12 times the monthly premium). The estimated risk premiums (in percent) are reported in the first line corresponding to each model. The second line contains the Fama–MacBeth standard errors (in parentheses and percent). The third line displays the Fama–MacBeth t-statistics (in square brackets). The levels of statistical significance of the estimated risk premiums are indicated by italics (10% level), an underline (5% level), and boldface (1% level). The second to last column shows the average absolute pricing errors (in percent), which we calculate as $\frac{1}{N}\sum_{i=1}^{N}\left|\frac{1}{T}\sum_{t=1}^{T}\hat{\alpha}_{i,t}\right|$, where $\hat{\alpha}_{i,t}$, $i = 1, \ldots, N$, denotes the pricing errors of the N industries in month $t = 1, \ldots, T$. The statistic MAE is annualized by multiplication by 12. The last column presents the cross-sectional R^2 value of each model, following Jagannathan and Wang (1996), Petkova (2006), and others. The sample period is August 1990 to December 2011 and 59 months are lost when the factor loadings are estimated by rolling regressions.

Panel A: Risk loadings from full-period (static) regression

	λ_L	λ_{TERM}	λ_{DEF}	λ_{DY}	λ_{RF}	λ_{PE}	λ_{VS}	λ_{CP}	MAE	R^2
PS	-0.21								3.65	0.00
	(0.64)									
	[-0.33]									
FF5		0.51	-0.26						3.54	0.11
		(0.50)	(0.19)							
		[1.01]	[-1.38]							
HL		-0.31	-0.13						2.70	0.04
		(0.49)	(0.21)							
		[-0.62]	[-0.61]							
P		-0.84	-0.35	3.05	0.88				2.41	0.21
		(0.56)	(0.20)	(15.86)	(0.51)					
		[-1.48]	[*-1.75*]	[0.19]	[*1.72*]					
CV		0.06				6.92	-7.75		2.66	0.36
		(0.51)				(3.90)	(14.65)			
		[0.11]				[*1.78*]	[-0.53]			
KLVN		-0.15						-1.58	2.67	0.11
		(0.48)						(1.04)		
		[-0.31]						[-1.51]		

Table 26 (continued)

Panel B: Risk loadings from 60-month rolling regressions										
	λ_L	λ_{TERM}	λ_{DEF}	λ_{DY}	λ_{RF}	λ_{PE}	λ_{VS}	λ_{CP}	MAE	R^2
PS	-0.66 (0.42) [-1.55]								3.80	0.01
FF5		-0.61 (0.37) [-1.63]	-0.21 (0.16) [-1.26]						3.56	0.04
HL		-0.51 (0.39) [-1.30]	0.03 (0.17) [0.15]						2.43	0.38
P		-0.73 (0.40) [*-1.84*]	0.03 (0.16) [0.18]	-4.37 (16.70) [-0.26]	0.49 (0.33) [1.47]				2.19	0.07
CV		-0.53 (0.42) [-1.25]				5.43 (4.21) [1.29]	-27.44 (11.91) [<u>-2.30</u>]		2.74	0.86
KLVN		-0.47 (0.38) [-1.24]						-1.16 (0.72) [-1.61]	2.53	0.19

Table 27 shows the annualized results of the Fama–MacBeth (1973) regressions with the estimated slopes on both traded and non-traded factors as the independent variables (Eqs. (4.19) and (4.20)). The average absolute pricing errors and cross-sectional R^2 values indicate that the CAPM is least able to explain the mean industry excess returns. Instead, with full-period slope estimates, C shows the highest explanatory power (followed by PS), while P and PS seem to perform best with rolling slope estimates.

The estimated risk premium for a given factor again depends on both the asset pricing model it is included in and the method that is chosen for the estimation of the risk loadings. First, I consider the traded factors. The estimates for the market risk premium are more than one percentage point higher than the sample mean of RM. In addition, using full-period slope estimates, the standard errors of the estimates are nearly equal to that of the historical average of RM (approximately 3.58% p.a.); thus, the market risk premium is now significantly different from zero within all models apart from the CAPM. However, with rolling slope estimates, these standard errors are approximately 4.2% p.a., making the market risk premium insignificant within all models. Furthermore, the deviations of the estimated premiums for SMB and HML from their sample means are substantially greater. The standard errors of the premium estimates are higher than the standard errors of the sample means within all models and for both full-period and rolling slope estimates and the estimates are statistically insignificant in all instances. With full-period slopes, the risk premium for UMD is now as high as 19.53% p.a. and significantly different from zero, although the standard error is approximately three percentage points higher than that of the sample mean. Second, one can see that the estimated premiums for the non-traded factors deviate at least somewhat from the estimates obtained from Eq. (4.18). Their standard errors are, overall, comparable to those from Eq. (4.18).

Table 27. Factor risk premiums from cross-sectional regressions.

This table reports the results from monthly Fama–MacBeth (1973) regressions with the excess returns on 35 European industries as left-hand side variables. The right-hand side variables are the factor risk loadings, which are estimated in one static multiple time-series regression (Panel A) or 60-month rolling multiple time-series regressions (Panel B) for each model. The regressions are estimated without an intercept term. The variables λ_L, λ_{TERM}, λ_{DEF}, λ_{DY}, λ_{RF}, λ_{PE}, λ_{VS}, and λ_{CP} represent the estimated risk premiums for the liquidity factor, slope of the yield curve, corporate bond default spread, market dividend-to-price ratio, short-term risk-free rate, aggregate price–earnings ratio, value spread, and Cochrane–Piazzesi factor. All risk premiums are annualized (12 times the monthly premium). The estimated risk premiums (in percent) are reported in the first line corresponding to each model. The second line contains the Fama–MacBeth standard errors (in parentheses and in percent). The third line displays the Fama–MacBeth t-statistics (in square brackets). The levels of statistical significance of the estimated risk premiums are indicated by italics (10% level), an underline (5% level), and boldface (1% level). The second to last column shows the average absolute pricing errors (in percent), which we calculate as $\frac{1}{N}\sum_{i=1}^{N}\left|\frac{1}{T}\sum_{t=1}^{T}\hat{\alpha}_{i,t}\right|$, where $\hat{\alpha}_{i,t}$, $i = 1, \ldots, N$, denotes the pricing errors of the N industries in month $t = 1, \ldots, T$. The statistic MAE is annualized by multiplication by 12. The last column presents the cross-sectional R^2 value of each model, following Jagannathan and Wang (1996), Petkova (2006), and others. The sample period is August 1990 to December 2011 and 59 months are lost when the factor loadings are estimated by rolling regressions.

Panel A: Risk loadings from full-period (static) regression

	λ_{RM}	λ_{SMB}	λ_{HML}	λ_{UMD}	λ_L	λ_{TERM}	λ_{DEF}	λ_{DY}	λ_{RF}	λ_{PE}	λ_{VS}	λ_{CP}	MAE	R^2
CAPM	5.80												2.68	-0.59
	(3.59)													
	[1.61]													
FF3	6.16	-3.16	0.07										2.55	-0.50
	(3.58)	(2.45)	(2.90)											
	[*1.72*]	[-1.29]	[0.02]											
C	7.31	-3.85	1.10	19.53									1.88	0.18
	(3.58)	(2.45)	(3.03)	(6.39)										
	[2.04]	[-1.57]	[0.36]	[**3.06**]										
PS	6.35	-2.11	-3.42		-2.13								1.98	0.16
	(3.58)	(2.48)	(2.77)		(0.59)									
	[*1.78*]	[-0.85]	[-1.24]		[**-3.62**]									
FF5	6.33	-2.61	-0.43			-0.43	-0.22						2.44	-0.42
	(3.58)	(2.28)	(2.97)			(0.51)	(0.19)							
	[*1.77*]	[-1.14]	[-0.14]			[-0.84]	[-1.13]							
HL	6.05					-0.47	-0.10						2.53	-0.50
	(3.58)					(0.49)	(0.21)							
	[*1.69*]					[-0.96]	[-0.46]							
P	6.69					-1.26	-0.38	5.84	1.21				2.04	-0.03
	(3.58)					(0.56)	(0.20)	(15.85)	(0.52)					
	[*1.87*]					[-2.23]	[*-1.88*]	[0.37]	[2.34]					
CV	5.93					-0.22				7.40	-4.01		2.54	-0.49
	(3.58)					(0.49)				(3.91)	(14.71)			
	[*1.66*]					[-0.45]				[*1.89*]	[-0.27]			
KLVN	6.07					-0.34						-1.70	2.60	-0.39
	(3.58)					(0.47)						(1.04)		
	[*1.69*]					[-0.72]						[-1.63]		

Table 27 (continued)

Panel B: Risk loadings from 60-month rolling regressions

	λ_{RM}	λ_{SMB}	λ_{HML}	λ_{UMD}	λ_{L}	λ_{TERM}	λ_{DEF}	λ_{DY}	λ_{RF}	λ_{PE}	λ_{VS}	λ_{CP}	MAE	R^2
CAPM	6.01												2.80	-0.45
	(4.24)													
	[1.42]													
FF3	5.59	0.09	-0.05										2.42	-0.08
	(4.19)	(2.56)	(2.75)											
	[1.33]	[0.04]	[-0.02]											
C	5.96	-0.97	-0.48	4.00									2.20	0.07
	(4.20)	(2.52)	(3.03)	(5.14)										
	[1.42]	[-0.38]	[-0.16]	[0.78]										
PS	5.99	-0.46	-0.59		-0.95								2.06	0.20
	(4.19)	(2.48)	(2.73)		(0.38)									
	[1.43]	[-0.18]	[-0.22]		[-2.53]									
FF5	5.99	-0.29	-0.40			-0.68	-0.09						2.14	-0.03
	(4.19)	(2.54)	(2.69)			(0.37)	(0.16)							
	[1.43]	[-0.11]	[-0.15]			[-*1.85*]	[-0.57]							
HL	6.68					-0.53	0.03						2.21	0.03
	(4.23)					(0.39)	(0.17)							
	[1.58]					[-1.37]	[0.16]							
P	6.57					-0.78	0.02	-4.34	0.55				1.98	0.12
	(4.15)					(0.39)	(0.16)	(16.55)	(0.32)					
	[1.59]					[-2.02]	[0.13]	[-0.26]	[*1.69*]					
CV	5.82					-0.50				5.14	-27.59		2.68	-0.26
	(4.20)					(0.41)				(4.24)	(11.82)			
	[1.38]					[-1.22]				[1.21]	[-2.33]			
KLVN	6.41					-0.46						-0.99	2.38	-0.18
	(4.22)					(0.38)						(0.71)		
	[1.52]					[-1.23]						[-1.39]		

Last but not least, the risk premiums estimated with cross-sectional regressions also depend on the set of test assets. They change when one considers, for instance, portfolios sorted by size and the book-to-market ratio or size and momentum instead of industry portfolios, as the GMM results in Sec. 5.2.2 show.

5.3.3. Industry costs of equity capital

Tables 28 and 29 show the results of the four different ways to estimate CEs described in Sec. 4.3.3. I first compare the different CE estimates obtained by these four approaches within a given asset pricing model. The global picture is that the average CEs across all industries, that is, the CEs of the typical industry, seem to be approximately one percentage point higher when all factor premiums are estimated with Fama–MacBeth regressions instead of sample means, which coincides with the higher market risk premium that I obtained from the Fama–MacBeth regressions. This difference is slightly higher for the rolling regression slope estimates, although the resulting deviations between the CE estimates of the typical industry using full-period instead of rolling slopes are rather small. Regarding the individual industries, all of these differences seem to be considerably greater. For instance, *Software & Computer Services* has a CE of -4.30% for C using factor sample means but a CE of 6.04% for C with Fama–MacBeth regressions, with full-period risk loadings in both instances.

In the next step, I compare the CE estimates across models, focusing on the CEs obtained with full-period slope estimates and risk premiums estimated by sample means and Eq. (4.18).

The CEs of the typical industry do not differ much across models: The highest value is obtained using PS (4.97%), whereas the lowest required rate of return is obtained using P (4.33%). However, once again, the differences for individual industries are much greater. For instance, the CE of *Life Insurance* ranges between 4.65% (according to C) and 11.29% (according to FF5), whereas the CAPM estimate is 6.23%.

Overall, the CE estimates of the typical industry do not differ much according to the approach used to estimate the risk loadings and risk premiums or according to the asset pricing model used. However, all of these differences are much greater for individual industries. Some deviations are even distressing and are probably even greater for individual firms, which confirms the overall results of FF97 and GM09.

Table 28. Industry CEs in excess of the risk-free rate, with factor risk premiums from sample means and cross-sectional regressions.
This table displays the estimated CE in excess of the risk-free rate for 35 European industries defined in Table 9. The last row shows their mean values across all 35 industries. The factor risk loadings are obtained from either full-period or 60-month rolling multiple time-series regressions, as conducted in Tables 21 and 22. The factor risk premiums associated with the traded factors are estimated by their sample means and are taken from Table 25. The risk premiums associated with the non-traded factors are estimated by Fama–MacBeth regressions conducted under the null hypothesis that the risk premiums associated with a model's traded factors are equal to their sample means, using factor risk loading estimates from either full-period or 60-month rolling multiple time-series regressions and are taken from Table 26. The results are annualized and in percent.

	CAPM		FF3		C		PS		FF5		HL		P		CV		KLVN	
	Full period	Rolling	Full period	Rolling	Full period	Rolling	Full period	Rolling	Full period	Rolling	Full period	Rolling	Full period	Rolling	Full period	Rolling	Full period	Rolling
Ind1	4.02	4.86	5.12	0.20	5.00	3.11	5.36	-2.38	5.94	-1.56	3.42	3.07	4.92	4.20	4.30	-2.38	4.18	2.62
Ind2	4.76	5.06	6.81	4.21	7.08	9.16	6.79	1.90	7.39	3.99	4.73	4.34	4.38	3.31	4.81	2.53	4.39	5.37
Ind3	4.75	5.35	7.97	10.38	5.93	7.42	7.91	11.23	7.81	8.51	3.40	7.37	0.88	6.74	4.80	12.98	3.39	5.64
Ind4	6.52	8.22	5.55	1.07	5.29	5.92	5.56	-0.85	8.92	1.16	5.86	5.01	5.83	4.60	6.12	-4.77	6.65	9.31
Ind5	5.43	7.00	5.11	-0.17	5.30	0.65	5.40	-2.30	8.28	-0.70	4.75	5.05	3.83	4.18	4.27	-4.22	5.21	10.69
Ind6	5.18	6.10	7.34	7.99	5.95	6.23	7.42	6.13	6.79	5.65	4.90	5.44	1.68	4.39	4.99	8.12	5.17	5.67
Ind7	4.83	4.04	5.65	0.73	4.18	-1.45	5.00	-1.21	4.63	0.92	4.69	2.95	5.06	3.97	5.76	4.21	4.79	4.47
Ind8	5.12	6.39	6.56	5.80	7.56	8.99	6.55	3.57	6.79	5.04	5.29	5.49	3.98	5.20	4.62	4.46	6.54	7.79
Ind9	6.64	5.98	2.54	8.11	0.37	9.62	2.14	8.10	3.68	8.19	6.04	5.16	5.60	3.64	6.02	5.21	4.49	6.72
Ind10	5.54	6.28	4.89	3.25	3.81	6.00	4.71	2.22	5.66	2.73	5.36	5.50	4.94	5.36	5.58	8.63	4.74	8.07
Ind11	4.68	4.95	4.93	4.64	3.92	2.85	5.17	5.55	3.42	5.04	5.22	5.20	3.74	5.27	5.33	8.15	4.56	6.79
Ind12	4.75	4.35	2.63	0.52	2.61	-1.08	2.75	-0.09	2.45	0.00	4.57	3.56	4.77	4.51	6.09	4.15	4.53	4.42
Ind13	5.54	4.90	6.83	6.26	6.15	10.08	7.18	6.20	4.55	3.91	4.34	6.09	2.79	5.37	4.53	-0.01	4.81	4.59
Ind14	3.01	3.83	4.60	-0.98	5.44	-2.63	5.16	1.48	5.09	0.49	3.26	3.04	4.10	3.43	3.71	5.33	2.80	3.85
Ind15	2.53	2.20	4.55	-1.43	6.19	-0.10	5.26	2.10	2.41	0.28	3.15	3.72	4.26	5.12	2.91	7.51	2.99	3.98
Ind16	4.11	3.61	4.03	-0.90	3.64	-2.84	4.34	-0.17	2.81	-0.17	4.43	3.67	5.64	5.10	4.56	7.28	4.07	3.42
Ind17	5.21	5.75	4.67	5.00	3.25	5.05	4.69	4.82	6.36	5.11	5.07	4.07	5.16	3.40	5.93	7.41	6.06	5.99
Ind18	3.99	4.72	7.88	5.19	9.30	4.66	8.24	6.69	7.44	5.65	4.55	6.19	3.86	5.67	4.92	7.67	5.76	7.25
Ind19	2.06	2.18	5.23	-3.83	7.29	-2.28	6.45	-1.30	3.67	-2.00	3.10	1.93	5.13	4.49	3.92	6.09	3.72	1.65
Ind20	3.46	2.79	1.93	-0.88	2.92	0.51	1.95	-0.80	1.48	0.41	3.32	3.50	3.66	4.68	3.40	6.06	4.21	4.73
Ind21	2.28	1.65	3.89	0.74	5.35	0.77	3.88	0.31	2.31	1.42	2.50	3.07	2.21	3.45	3.84	2.92	4.49	3.34
Ind22	3.25	3.17	3.74	0.48	3.18	0.56	3.64	0.77	3.00	1.30	3.40	3.08	4.39	4.09	4.31	7.46	3.60	2.53
Ind23	4.07	4.29	4.47	1.43	2.56	-1.68	4.64	2.90	4.38	2.37	4.30	4.01	2.28	4.64	4.31	5.53	3.90	4.64
Ind24	5.25	4.47	-0.55	5.16	0.29	5.78	-1.31	3.81	-0.22	4.27	4.87	4.18	5.15	4.54	3.65	7.98	4.80	4.75
Ind25	5.04	4.75	2.46	3.86	2.36	2.55	2.07	1.69	0.94	3.35	5.04	4.99	4.90	5.86	4.34	6.68	4.20	5.46
Ind26	4.62	2.57	1.11	3.81	2.73	7.34	0.75	2.84	2.53	5.12	4.56	3.14	6.34	4.08	3.33	4.61	4.48	2.63
Ind27	3.18	4.14	4.32	5.14	5.72	8.29	4.40	2.12	3.01	4.61	3.84	3.76	3.81	4.00	4.05	3.39	3.72	3.18
Ind28	3.39	3.71	7.36	8.51	7.89	9.83	7.61	7.44	5.71	8.21	4.00	4.15	3.93	3.57	3.05	3.95	3.46	1.45
Ind29	5.96	7.31	10.83	15.75	7.58	6.15	11.06	20.27	10.68	16.39	6.55	8.89	4.33	6.99	7.28	10.46	5.44	6.68
Ind30	5.56	4.99	10.52	13.31	7.00	9.17	10.67	14.87	8.98	12.97	5.63	6.12	1.76	4.60	6.17	8.67	5.45	3.78
Ind31	6.23	7.54	10.70	14.89	4.65	5.83	10.24	18.11	11.29	16.00	7.23	7.64	4.82	6.51	7.26	9.88	5.05	6.07
Ind32	3.44	4.68	5.25	6.00	4.70	4.66	5.66	8.71	4.17	7.84	4.46	5.22	3.96	4.55	4.84	10.63	4.40	6.52
Ind33	4.52	4.81	4.33	3.39	4.20	0.70	4.63	4.99	3.11	3.76	4.82	4.95	4.77	4.99	5.02	6.47	5.12	5.29
Ind34	6.53	4.39	-3.28	2.88	-4.30	1.28	-3.33	0.76	-0.98	3.25	5.48	2.90	7.93	3.05	4.40	3.63	2.80	2.09
Ind35	7.81	5.51	2.20	6.15	-0.19	-0.79	1.47	5.16	3.79	4.67	6.45	4.98	6.72	3.90	7.65	6.19	9.22	5.89
Means	4.66	4.76	4.92	4.19	4.43	3.72	4.97	4.16	4.81	4.23	4.64	4.61	4.33	4.61	4.86	5.51	4.66	5.07

Table 29. Industry CEs in excess of the risk-free rate, with factor risk premiums from cross-sectional regressions.
This table displays the estimated CE in excess of the risk-free rate for 35 European industries defined in Table 9. The last row shows their mean values across all 35 industries. The factor risk loadings are obtained from either full-period or 60-month rolling multiple time-series regressions, as conducted in Tables 21 and 22. The factor risk premiums are estimated by Fama–MacBeth regressions using factor risk loading estimates from either full-period or 60-month rolling multiple time-series regressions and are taken from Table 27. The results are annualized and in percent.

	CAPM		FF3		C		PS		FF5		HL		P		CV		KLVN	
	Full period	Rolling	Full period	Rolling	Full period	Rolling	Full period	Rolling	Full period	Rolling	Full period	Rolling	Full period	Rolling	Full period	Rolling	Full period	Rolling
Ind1	4.86	6.08	4.83	6.75	5.77	7.97	5.97	3.76	3.99	5.08	4.12	4.91	6.52	6.91	4.59	-1.23	4.85	4.28
Ind2	5.75	6.33	6.26	6.31	8.06	8.07	5.47	3.63	6.23	5.95	5.84	6.26	5.88	4.74	5.78	3.74	5.56	6.90
Ind3	5.74	6.69	5.45	4.82	3.97	4.11	3.48	5.96	2.88	5.98	4.22	9.88	1.18	8.53	4.90	15.90	4.05	7.58
Ind4	7.88	10.28	6.01	9.90	7.11	10.55	5.73	7.33	5.98	8.38	6.25	7.65	6.42	7.32	6.29	-5.50	7.27	11.53
Ind5	6.56	8.76	5.97	9.97	7.52	11.11	7.80	8.09	5.77	9.12	5.24	7.61	4.24	6.46	4.78	-3.79	5.81	12.17
Ind6	6.25	7.63	5.70	6.25	5.17	5.81	5.05	3.63	4.99	4.89	6.00	7.90	2.25	6.25	5.83	11.34	6.30	7.63
Ind7	5.84	5.05	5.63	4.84	4.71	4.16	0.87	2.15	5.21	4.63	5.94	4.30	7.09	5.91	6.50	5.16	6.03	5.65
Ind8	6.18	7.99	6.19	7.11	9.10	7.86	5.49	4.12	6.51	6.51	6.47	7.87	5.30	7.41	5.90	4.65	7.82	9.58
Ind9	8.03	7.48	7.87	6.45	5.68	7.14	6.98	6.76	7.87	5.86	7.40	7.34	7.40	4.98	7.36	6.73	5.85	8.52
Ind10	6.69	7.85	5.77	7.34	5.43	7.88	4.44	6.10	5.89	7.31	6.29	7.84	6.16	7.57	6.25	9.88	5.80	9.80
Ind11	5.66	6.19	5.00	5.40	4.67	4.94	6.23	6.81	5.66	6.35	6.51	7.09	5.21	7.00	6.40	8.71	5.98	8.23
Ind12	5.74	5.44	5.12	5.14	5.96	4.46	6.62	4.31	5.19	5.07	5.60	5.14	6.46	6.58	6.50	5.28	5.60	5.76
Ind13	6.69	6.13	6.97	6.49	7.44	8.77	8.83	7.52	4.45	6.75	5.95	8.52	4.60	7.32	6.01	4.28	6.14	6.28
Ind14	3.63	4.79	3.64	5.27	5.83	5.08	6.66	9.26	4.07	6.04	3.89	4.27	5.32	4.85	4.03	5.95	3.55	5.09
Ind15	3.06	2.76	3.56	3.54	6.87	4.29	7.62	9.18	3.95	6.47	4.32	4.68	6.41	6.53	3.99	8.14	4.18	4.68
Ind16	4.96	4.51	4.43	4.79	4.84	4.34	6.28	6.11	4.84	6.10	5.54	5.00	7.75	7.08	5.46	7.40	5.26	4.70
Ind17	6.30	7.19	5.72	6.41	4.78	6.43	5.84	6.35	6.11	5.35	5.86	6.01	6.36	5.06	6.30	7.51	6.99	7.69
Ind18	4.82	5.90	6.26	5.80	9.80	6.43	7.56	8.63	6.75	7.79	5.96	8.25	5.70	7.12	6.11	8.62	7.20	8.62
Ind19	2.49	2.72	3.29	4.03	7.25	4.88	10.40	8.36	4.33	5.46	4.16	2.62	7.53	6.53	4.49	6.99	4.87	2.48
Ind20	4.18	3.49	3.90	3.71	6.02	4.00	4.64	4.04	3.88	5.76	4.18	4.56	5.07	6.16	4.15	6.24	5.06	5.51
Ind21	2.76	2.06	4.10	2.82	7.00	3.68	3.90	2.90	3.99	4.72	3.68	3.94	3.90	4.27	4.50	2.13	5.57	3.80
Ind22	3.93	3.96	4.36	4.27	4.35	4.59	3.70	5.12	4.49	5.13	4.43	4.23	6.30	5.72	4.96	8.27	4.64	3.70
Ind23	4.92	5.36	5.18	5.28	3.34	4.64	6.29	7.68	5.62	6.17	5.39	5.54	3.26	6.34	5.23	6.34	5.05	6.03
Ind24	6.34	5.59	5.96	4.96	7.73	5.32	3.35	3.27	6.31	4.57	6.04	5.96	7.11	6.43	5.17	8.79	5.97	6.14
Ind25	6.09	5.93	4.91	5.08	5.67	4.52	3.01	1.93	5.05	5.72	6.27	6.86	6.77	8.03	5.69	8.60	5.48	6.97
Ind26	5.58	3.21	6.95	3.31	9.98	4.85	6.81	2.49	7.80	4.42	5.84	4.15	9.07	5.42	4.99	3.88	5.75	3.58
Ind27	3.84	5.18	3.70	4.59	6.71	5.57	3.65	0.51	4.46	4.19	4.87	5.37	5.36	5.72	4.86	4.39	4.85	4.65
Ind28	4.10	4.64	5.13	3.76	7.21	4.70	5.44	2.59	5.38	3.71	5.38	5.71	5.94	4.73	4.62	6.84	4.86	3.08
Ind29	7.21	9.14	7.97	6.50	5.14	4.50	7.66	13.00	8.48	7.99	8.18	11.84	5.91	8.50	8.42	13.71	7.21	9.47
Ind30	6.71	6.24	8.21	4.19	4.89	3.60	7.66	6.49	7.55	4.55	7.53	8.23	3.21	5.60	7.76	9.43	7.28	5.78
Ind31	7.53	9.43	8.38	6.57	1.38	4.30	3.70	10.97	9.83	6.97	8.89	10.38	6.36	8.37	8.67	10.97	6.96	8.92
Ind32	4.15	5.85	3.17	4.20	3.44	3.25	4.67	7.72	4.51	5.95	5.33	6.83	5.09	5.53	5.33	9.74	5.51	7.99
Ind33	5.47	6.02	4.86	5.62	5.74	5.11	6.65	8.19	5.24	6.55	6.02	6.80	6.59	6.70	5.96	6.56	6.40	6.87
Ind34	7.89	5.49	7.27	5.16	6.04	4.88	11.15	2.34	7.51	4.36	6.54	4.33	10.44	4.68	5.91	3.38	3.84	3.76
Ind35	9.44	6.89	10.90	6.34	8.41	5.23	9.23	5.38	9.97	5.50	8.35	7.19	9.69	5.50	9.17	6.63	10.88	7.55
Means	5.64	5.95	5.67	5.51	6.09	5.63	5.97	5.79	5.74	5.87	5.79	6.43	5.94	6.34	5.80	6.45	5.84	6.60

6. Discussion

6.1. Comparing the Results with Prior Empirical Evidence

Research Question 1 asks whether aggregate returns on commodity futures are predictable.[39] The results presented in Sec. 5.1 indicate that this question deserves an affirmative answer. They show that many variables exhibit predictive power over aggregate commodity futures returns in-sample (IS). Moreover, they indicate that aggregate returns on commodity futures are predictable out-of-sample (OOS), especially by the combination forecasts proposed by Rapach et al. (2010). Hence, it seems that expected returns on the commodity futures market are not constant but vary through time. These results confirm many of the findings that are documented in the literature on commodity returns presented in Sec. 2.1.2 within a different data set. Accordingly, I demonstrate that US monetary policy (measured by *M2* in this thesis), bond yields (*RF* and *LTY*), the inflation rate (*INFL*), the term structure spread (*TERM*), and the default spread (*DEF* and *DFR*), as well as other macroeconomic variables, have predictive power over returns on commodity futures. Moreover, the results confirm that several commodity price measures show forecasting power (especially the commodity variance *CVAR* and commodity spot momentum *C12CM_spot* in my analysis). Furthermore, the results verify the findings of Moskowitz et al. (2012) and show that there is significant time-series momentum (particularly within the IS period January 1972 to December 1999 and the OOS period January 1980 to December 1999). Additionally, the outcomes identify several significant IS predictors that have not yet been considered for commodity futures. For instance, stock market investor sentiment (*SENT*), stock market liquidity (*CL*), the consumption–wealth ratio (*CAY*), and the cumulative equity premium (*CRMRF*) all seem to predict commodity futures market returns reasonably well.

Beyond that, the outcomes of Research Question 1 point to some commonalities, but also to a variety of differences in the individual factors that predict commodity futures and US stock returns. In the following, I outline some key commonalities and differences regarding the IS single predictive regressions over the full sample period January 1972 to June 2010 (Panel A of Table 10, p. 85). It seems that there are only a few commonalities: For instance, the results for *SENT* and the dividend–payout ratio (*D/E*) coincide with the finding of Baker and Wurgler (2007) that a high level of *SENT* forecasts low US stock market returns and the evidence of Lamont (1998) that *D/E* predicts rising returns on US stocks. Moreover, the finding that *CVAR* significantly forecasts rising long-term (48- and 60-month) aggregate returns on commodity futures agrees with Guo's (2006a) result that stock variance (*SVAR*), which is the stock analogy of *CVAR*, predicts rising US stock returns. Beyond that, the results in Table 12 (p. 93) validate the finding of Rapach et al. (2010) that US stock returns are predictable OOS when applying combination instead of individual forecasts.

However, there seem to be a variety of differences: For example, I observe that none of the results of Maio and Santa-Clara (2012, hereafter MSC), who conduct predictive regressions

[39] This chapter is partly based on Sec. 5 of Lutzenberger (2014a).

over the period July 1963 to December 2008 with the US stock market return as the dependent variable and seven of the factors I employ for commodity futures as the independent variable, coincide with the results in Panel A of Table 10, since the predictive slope on each factor differs in either its sign or its statistical significance. In particular, there are factors that are significant for both asset classes but whose predictive slopes show opposite signs. These predictors are the stock market dividend–price ratio (*DY*), the Treasury bill rate (*RF*), and *CL*, whose predictive slopes seem to be significantly negative for commodity futures, but, according to MSC, significantly positive for US stocks. Similarly, the stock market earnings–price ratio (*E/P*) shows a negative slope in Panel A of Table 10, although the predictive slope on the price–earnings ratio (the reciprocal of *E/P*), is negative for MSC. Moreover, the default yield spread (*DEF*) and the Cochrane–Piazzesi factor (*CP*) seem to have no forecasting power for commodity futures, at least not over the period January 1972 to June 2010, while these two variables predict US stock returns significantly positively according to MSC. Furthermore, *TERM* seems to be a significant predictor of rising commodity futures returns, but, as shown by MSC, it is insignificant for US stocks. Beyond that, *CAY* appears to predict commodity futures returns significantly negatively, while this factor is, according to Lettau and Ludvigson (2001a), a positive predictor of US stock market returns. Finally, Welch and Goyal (2008) conclude that the long-term yield (*LTY*) and the default return spread (*DFR*) are not significant IS predictors of S&P 500 returns at the annual frequency, although the results in Panel A of Table 10 indicate that both variables are significant predictors of aggregate returns on commodity futures one year ahead.

Research Question 2 asks whether the CAPM and Fama and French's (1993) three-factor model (FF3), Carhart's (1997) four-factor model (C), Pástor and Stambaugh's (2003) four-factor model (PS), Fama and French's (1993) five-factor model (FF5), Campbell and Vuolteenaho's (2004) unrestricted version of the ICAPM (CV), Hahn and Lee's (2006) ICAPM (HL), Petkova's (2006) ICAPM (P), and the three-factor model of Koijen et al. (2010; KLVN) are able to capture the cross section of average returns on European stocks. Some of the first-stage generalized method of moments (GMM) results displayed in Tables 17 (p. 105) and 18 (p. 107) counter the prior evidence from the US stock market that is reported in MSC's Tables 7 and 8. In particular, both HL and KLVN appear to be unable to explain the cross section of expected stock returns in Europe, while they seem to capture the variation in average returns on US stocks, regardless of whether they are tested over the 25 size/book-to-market portfolios (SBM25) or the 25 size/momentum portfolios (SM25). Moreover, the CV model shows low explanatory power for the SM25 in Europe. Nevertheless, many facts confirm the US evidence: The CAPM seems to be unable to capture the average returns on SBM25 and SM25, both FF3 and PS appear to explain the value effect reflected by SBM25 but have difficulties explaining the momentum effect represented by SM25, CV shows good performance when tested with SBM25, and P, C, and FF5 seem to explain both sets of testing assets reasonably well. All in all, C seems to be most qualified to describe the cross section of expected European stock returns, which confirms MSC's US evidence.

Additionally, I compare the results of Research Question 2 to evidence documented in prior European studies. Four of the nine asset pricing models I investigate have already been tested on European data: the CAPM, FF3, C, and FF5. Fama and French (2012) investigate, among other things, the CAPM, FF3, and C over 25 size/book-to-market portfolios of European stocks using time-series regressions over the sample period November 1990 to March 2011.[40] The respective results are reported in Table 3 of their paper. They find that the CAPM is rejected by the test of Gibbons, Ross, and Shanken (1989, hereafter GRS) at the 5% level, while neither FF3 nor C can be rejected.[41] These results confirm my conclusion based on the GMM results reported in Tables 17 (p. 105) and 18 (p. 107) that both FF3 and C are able to capture the average returns on SBM25, while the CAPM is unable to explain these portfolios. Moreover, the authors show that the average absolute intercepts in these time-series regressions (which correspond to the average absolute pricing errors) are 0.20% for the CAPM, 0.09% for FF3, and 0.07% for C. These numbers are somewhat smaller than the mean absolute pricing errors that I report for the GMM estimation in Panel A of Table 17 and in Panel A of Table 18, which are 0.24%, 0.13%, and 0.10% for the CAPM, FF3, and C, respectively. Nevertheless, their ranking order and their implications are similar. Furthermore, Fama and French test the C model over 25 size/momentum portfolios of European stocks in a similar way and document the respective results in Table 6 of their paper.[42] The GRS test strongly rejects C and its average absolute time-series regression intercept is 0.18%. While this number is quite similar (although higher) to the mean absolute pricing error of 0.16% that I report for C in Panel B of Table 18, the model's rejection by the GRS test somewhat counters my conclusion that C is able to explain the SM25 portfolio returns reasonably well. In fact, Fama and French (2012, p. 470) conclude that their European results associated with the C model and size/momentum portfolio returns are "disappointing."

Bauer et al. (2010) find that FF3 is strongly rejected by the GRS test over 25 European stock portfolios sorted on size and the book-to-market ratio (the p-value of the GRS statistic is 0.0003; see Bauer et al., 2010, Table 3). Hence, their outcomes counter my results of Research Question 2 to some extent, which might be due to the different sample period they consider (February 1985 to June 2002). Wallmeier and Tauscher (2014) examine both the CAPM and FF3 over 25 portfolios of European stocks sorted on size and the book-to-market ratio, conducting time-series regressions. Their sample period is July 1990 to December 2009. They report in Table 3 of their paper that seven of the time-series intercepts associated with the CAPM are significantly different from zero at the 5% level, while all 25 FF3 intercepts seem to be statistically equal to zero. I additionally compute their average absolute intercepts and obtain 0.23% for the CAPM and 0.08% for FF3. These values are smaller than the mean absolute pricing errors I report. Furthermore, FF3 cannot be rejected by the GRS test.[43] These results confirm my conclusion that the CAPM is unable to explain the average returns of SBM25, while FF3 does a reasonable job of capturing these portfolio returns.

To the best of my knowledge, no further studies investigate these asset pricing models on similar data for the whole European stock market. Nevertheless, a variety of studies employ

[40] The authors employ both global and local factors. To compare their outcomes with mine, I consider the results associated with the local factors that are constructed using exclusively European stocks.

[41] The GRS test investigates whether the null hypothesis that the intercepts in the time-series regressions associated with all 25 portfolios are jointly zero can be rejected.

[42] They do not report results for the CAPM and FF3 tested over size/momentum portfolios, referring to earlier international evidence that these models are unable to explain momentum.

[43] The authors do not report a GRS statistic for the CAPM.

data from individual European countries (see also Sec. 2.2.2). One of these studies, that of Ziegler et al. (2007), investigates FF5 on 16 portfolios of German stocks sorted by size and the book-to-market ratio using time-series regressions, considering the sample period July 1968 to June 1995. None of the intercepts associated with the 16 FF5 time-series regressions is significantly different from zero (reported in Table 6 of Ziegler et al., 2007). When I compute the average absolute intercept, I obtain a value of 0.07%, which is considerably smaller than the mean absolute pricing error associated with the GMM estimation of FF5 reported in Panel A of Table 18 (p. 107, 0.14%).[44] Hence, my conclusion that FF5 is able to capture the SBM25 portfolio returns is confirmed by Ziegler et al. for the German stock market. Finally, I consider the studies of Artmann et al. (2012a, 2012b) to further challenge my results associated with the SM25 portfolio returns (Panel B of Table 17, p. 105, and Panel B of Table 18, p. 107). Artmann et al. (2012a) conduct Fama–MacBeth (1973) regressions over the sample period 1963 to 2006 using 16 portfolios of German stocks sorted by size and momentum and report cross-sectional R^2 values of 0.35 for the CAPM, 0.69 for FF3, and 0.91 for C in Table 8 of their paper. Hence, their results confirm my conclusion that C does a considerably better job explaining the SM25 portfolio returns than the CAPM and FF3. Artmann et al. (2012b) document the results of time-series regressions for the CAPM, FF3, and C over 16 portfolios of German stocks sorted by size and momentum in Table 8 of their paper. They show that 11 of the 16 CAPM intercepts, 11 of the 16 FF3 intercepts, but only six of the 16 intercepts associated with the C model are significantly different from zero at the 10% level. When I calculate the average absolute of these intercepts, I obtain 0.37% for the CAPM, 0.33% for FF3, and 0.13% for C. Consequently, the pricing errors associated with C seem to be considerably lower than those of the CAPM and FF3. Moreover, Artmann et al. (2012b) find that both the CAPM and FF3 are strongly rejected by the GRS test at the 1% level, while C can only be rejected at the 5% level. These results confirm my conclusion that both the CAPM and FF3 are unable to capture the SM25 portfolio returns, while C does better in describing these portfolios. All in all, my results for Research Question 2 appear to be, with some exceptions, confirmed by prior European evidence.

Research Question 3 asks whether the multifactor models FF3, C, PS, FF5, CV, HL, P, and KLVN are consistent with the Intertemporal CAPM (ICAPM) within the European stock market. The outcomes presented in Table 19 (p. 108) disagree to some extent with MSC's prior US evidence (see MSC's Table 1): Both HL and KLVN seem to be consistent with the ICAPM only in Europe. Thereby, the HL model is ICAPM consistent when it is tested over SM25 and when one considers changes in investment opportunities driven by expected market returns (ICAPM criterion 3a), while the KLVN model is ICAPM consistent when the testing assets are SBM25 and when shifts in investment opportunities are driven by market volatility (ICAPM criterion 3b). In contrast, FF3 seems to be consistent with the ICAPM only in the US, where it is ICAPM consistent when it is estimated over SBM25 and regardless of whether one considers changes in investment opportunities driven by expected market returns or market volatility. The sole model that seems to be consistent in both the US and Europe is C. However, while C is ICAPM consistent in the US only when shifts in investment

[44] These values are, of course, not directly comparable to each other because of the different stock markets, the number of test assets, and the sample periods considered.

opportunities are driven by expected market returns (with both sets of testing assets), the model fulfills ICAPM criterion 3b in Europe (over both SBM25 and SM25) but does not meet criterion 3a when it is tested over the SM25. All remaining models seem to be inconsistent with the ICAPM in both the US and European stock markets.[45] Hence, the results of Research Question 3 indicate that MSC's (p. 610) following conclusion is not completely valid OOS:

> Overall, the Fama and French (1993) three-factor model performs the best in consistently meeting the ICAPM restrictions when investment opportunities are driven by the first two moments of aggregate returns when tested with the SBM25 portfolios. Apart from this model and the Carhart (1997) model, the other models cannot be justified with the ICAPM theory. The ICAPM is not really a "fishing license" after all.

In particular, it seems that C shows the best performance in meeting the ICAPM restrictions in Europe, while it looks as if FF3 is inconsistent with the ICAPM in this stock market. Nevertheless, the conclusion that the ICAPM cannot be termed a fishing license seems to be robust across different stock markets. The finding that C is most likely consistent with the ICAPM is remarkable, since quite a few of the studies presented in Sec. 2.2.1.4 suggest that momentum, which is represented by C's fourth factor, is of a behavioral nature and probably represents market inefficiencies. My results—as well as, to some extent, the US results of MSC—speak against this supposition by indicating that momentum is a systematic risk factor that captures the risk of shifts in an investor's investment opportunity set in the spirit of the ICAPM. Hence, these results suppose that the fact observed by Cochrane (2005, Sec. 20.2, p. 447), "that nobody wants to add [the momentum factor] as a risk factor," might be unjustified.

Research Question 4 asks whether the CAPM and the multifactor FF3, C, PS, FF5, CV, HL, P, and KLVN provide precise cost of equity capital (CE) estimates for European industries. The outcomes presented in Sec. 5.3 are twofold. First, they suggest that CE estimates for European industries that are obtained from the CAPM, FF3, and C are inaccurate and thus confirm prior evidence on these three models from US and UK industries provided by Fama and French (1997, hereafter FF97) and Gregory and Michou (2009, hereafter GM09) (see Sec. 2.3.3). Second, the six remaining multifactor models (FF5, PS, CV, HL, P, and KLVN) seem to do more harm than good, because their CEs appear to be even more imprecise than the CEs obtained from the CAPM, FF3, and C.

Research Question 5 asks whether the CE estimates for European industries, obtained from the nine asset pricing models I investigate throughout this thesis and different estimation techniques, differ from each other. The results displayed in Tables 28 (p. 130) and 29 (p. 131) indicate that the magnitudes of CE estimates for the typical industry do not differ much across the asset pricing models and estimation techniques I examine. However, these differences are much greater for individual industries. Some deviations are even distressing and are probably even greater for individual firms. Consequently, the choice of factor model seems to be quite important. This confirms the results of FF97, who report in Table 7 of their paper

[45] In contrast to the analysis of Research Question 2, there are no European studies with which my results can be compared.

considerable differences in the CEs obtained from the CAPM and FF3 for some industries. For instance, they document that the CE (i.e., the premium over the risk-free rate) of *Autos* is 5.13% p.a. when it is estimated using the CAPM, but 9.39% p.a. using FF3, with risk loadings from full-period regressions in both cases. For comparison, I report in Table 28 (p. 130) that the CAPM CE for *Automobiles & Parts* (Ind. 13) is 5.54% p.a. and the FF3 CE is 6.83% p.a., while P claims that the appropriate CE for this industry should be only 2.79% p.a. (employing risk loadings from full-period regressions in each case).[46]

6.2. Is the US Evidence Spurious?

I revisit the central question that motivates this thesis: Are the empirical facts that drive the second revolution in asset pricing, that is, time-varying expected market returns and multiple factors in the cross section of expected returns, real? Or are these observations specific to the limited sample of US stocks? On the one hand, it seems that these facts also apply to commodity futures and European stocks, respectively. In particular, my results suggest that expected returns on the market of commodity futures vary considerably through time and appear to support the view that time-varying expected returns are "a pervasive phenomenon" (Cochrane, 2011, p. 1051). Moreover, it looks as if multifactor models provide a better description of the cross section of average returns on European stocks than the CAPM does. On the other hand, there appear to be a variety of differences in the individual factors that drive expected commodity futures returns through time and expected European stock returns in the cross section. Moreover, it looks as if there are differences in the consistency of multifactor models with the ICAPM between the US and European stock markets. How can the differences between prior evidence from US stocks and my results for commodity futures and European stocks be interpreted?

These differences can be interpreted in at least two ways. On the one hand, one may assume that the same true process generates returns on different stock markets and asset classes (with or without time variation in expected returns), that is, the same set of factors drives both US stock returns and returns on other asset classes in the time series and one (or the same) set of common factors drives them in the cross section. Under this assumption, one can argue that the multitude of differing results with regard to the individual factors that drive expected returns suggests that the US evidence does not reflect the true return-generating process. Instead, one may conclude that the US evidence is a phenomenon that is specific to the limited sample of US stocks and caused by mere random chance. Therefore, my results might argue against the reliability of the US evidence that drives the second revolution in asset pricing.

On the other hand, it seems to be reasonable to assume that differing asset characteristics, institutional settings, and investor preferences result in return-generating processes that diverge across stock markets and asset classes, although their expected returns all vary

[46] A comparison of FF97's CEs for US industries and the CEs for European industries estimated here is, of course, difficult, since the risk profiles of a given industry probably differ in these two markets and both studies use quite different sample periods, that is, July 1963 to December 1994 versus August 1990 to December 2011.

through time and although all of them require multiple factors to explain their cross-sectional variations. Since different asset classes have different characteristics and dissimilar investors with different preferences (e.g., hedgers and speculators in futures markets vs. investors in stock markets) and because they are traded on different markets or exchanges with different institutional settings, this assumption is probably more realistic. For instance, the futures markets are zero-sum games so that expected futures returns or risk premiums may be both negative and positive (Szymanowska et al., 2014). Under this assumption, the differences with regard to the individual return-predicting factors I observe do not speak against prior evidence from the US; rather, they support the views that emerge from the US evidence.

6.3. Implications for Practice

The results of Research Question 1 suggest that portfolio managers, decision makers, and regulatory authorities should consider different factors in evaluating the state of the commodity futures market, although they also point to many common factors that drive expected returns on both commodities and US stocks. The presence of several common factors whose predictive slopes show opposite signs, such as the stock market dividend–price ratio (*DY*) and the Treasury bill rate (*RF*), indicates that expected returns on commodity futures are high in times when expected returns on US stocks are low. Hence, they suggest that commodity futures provide investors with a hedge against changes in expected US stock returns. Thus, my results contribute to studies such as that of Belousova and Dorfleitner (2012), who investigate the diversification benefits of commodities to a portfolio of traditional assets.

The outcomes of Research Questions 2 to 5 taken as a whole indicate that European decision makers and regulatory authorities should employ C to estimate CEs for several reasons if they insist upon using a (multi-)factor asset pricing model for this purpose.[47] First, this model appears to exhibit the greatest explanatory power for the cross section of expected stock returns in Europe. Second, this model seems to be most likely consistent with the ICAPM in the European stock market. Hence, it looks as if the empirically motivated C can be theoretically justified as a rational model that captures risk that is not considered by the CAPM, that is, the risk of changes in investment opportunities, rather than the irrational behavior of investors or market inefficiencies. Third, CEs estimated with C appear to be less imprecise than the CEs obtained from six other multifactor models I investigate (FF5, PS, CV, HL, P, and KLVN). Fourth, it looks as if one cannot simply employ another asset pricing model instead, such as the CAPM (which is less complex, since it comprises only one factor), since different models seem to result in quite different CEs. The finding that multifactor models that consider liquidity risk or that are explicitly justified by their authors as empirical applications of the ICAPM (PS, CV, HL, and P) appear to be unqualified in providing accurate CEs is quite unfortunate, since these multifactor models appear to have—prior to knowing the results of MSC and this thesis—considerable advantages over the CAPM, FF3, and C (see Ch. 1).

[47] I focus on the application of estimating CEs, although this conclusion is to some extent transferable to other applications, such as performance evaluation.

Nevertheless, the outcomes of FF97, GM09, and this thesis still suppose that using a (multi-)factor asset pricing model to estimate CEs at all is highly questionable. Even though the CE estimates of C appear to be less imprecise than those obtained from many other multifactor models, they are still *imprecise*. These results suppose that we might have to use alternative approaches to estimate CEs. For instance, they appear to confirm Cochrane (2011), who does not see a solution to practical problems such as project valuation in multifactor models but advocates the use of simple “comparables,” such as average returns on similar securities. Maybe the author is correct and we should, when estimating CEs, take one step backward from the CAPM to simpler models rather than forward to multifactor models. Or we should look for more innovative ways to obtain CEs, such as estimating the implied CE, as proposed by Gebhardt et al. (2001), among many others.

6.4. A Few Limitations

I should also discuss some limitations of my analysis that might affect the inferences I make. First, the research approach I follow to examine expected commodity futures returns in the time series, which is the investigation of a large set of potential predictors that are not a priori justified by any theory using standard predictive regressions, should be taken with a pinch of salt. To describe my concerns, I would like to cite an excerpt from the conclusion of Novy-Marx (2014, pp. 143–144):

> Standard predictive regressions fail to reject the hypothesis that the party of the US President, the weather in Manhattan, global warming, El Niño, sunspots, or the conjunctions of the planets are significantly related to anomaly performance [that is, they significantly predict future returns on a variety of anomaly strategies such as size, value, and momentum]. These results are striking and surprising. In fact, some readers might be inclined to reject some of this paper’s conclusions solely on the grounds of plausibility. I urge readers to consider this option carefully, however, as doing do so entails rejecting the standard methodology on which the return predictability literature is built.

While I employ potential predictors whose relation to expected returns appears to be more plausible than those of Novy-Marx’s, a critic may still argue that running large sets of standard predictive regressions without any theoretical basis for them probably leads to spurious inferences. Such a critic does, however, also reject the methodology that is standard in the predictability of returns literature.

Second, the two sets of testing assets I employ to investigate expected returns in the cross section of European stocks (SBM25 and SM25) represent only three return patterns observed in prior empirical studies that the CAPM seems to be unable to explain, namely, size, value, and momentum effects (Secs. 2.2.1.3 and 2.2.1.4). Both FF3 and C are empirical asset pricing models that are constructed to capture exactly these three effects, so that they are naturally playing home games when they are tested over SBM25 and SM25 (Fama and French, 2012). In particular, FF3 seeks to describe size and value and C seeks to capture size, value, and momentum. This is emphasized by Fama and French (2013, p. 4):

> The FF three-factor model is designed to capture the relation between average return and *Size* (market capitalization, price times shares outstanding) and the relation between average return and price ratios like the book-to-market ratio, which were the two well-known patterns in average returns at the time of our 1993 paper.

The two models are, however, not designed to capture other return patterns or CAPM anomalies such as profitability (Novy-Marx, 2013) and investment (Aharoni et al., 2013).[48] The same applies to C, with the exception of momentum. Consequently, the conclusion that C performs best in explaining the cross section of average returns in Europe is too general. A more appropriate conclusion is that the model does the best job in explaining the patterns in average returns on size, value, and momentum strategies for European stocks.

To shed further light on this issue, I consider the results of the Fama–MacBeth (1973) regressions displayed in Table 27 (p. 127) that are conducted using the 35 European industry return series, in addition to the results associated with the SBM25 and SM25 returns. In fact, Lewellen et al. (2010) advocate the use of industry portfolios instead of portfolios sorted on characteristics such as size and value in order to improve tests of asset pricing models. Note in Panel A of Table 27 that C, together with PS, appears to do the best job in capturing the dispersion in the average industry returns when factor loadings are estimated using full-period regressions. In particular, these two models show the lowest mean absolute pricing errors and the highest cross-sectional R^2 values. Moreover, C's momentum factor and the liquidity factor of PS are the sole factors that are significantly priced at the 1% level. Hence, C's success seems to be quite robust to using industry returns instead of returns on portfolios sorted by size, value, and momentum. Moreover, observe in Panel B of Table 27 that if we use 60-month rolling regressions to estimate factor loadings, C performs slightly worse and the momentum factor is no longer significantly priced. The best-performing models are now PS, whose liquidity factor is still significantly priced (at the 5% level), and P. Nevertheless, C still performs considerably better than both the CAPM and FF3 in capturing the industry returns.

Additionally, I check whether C is consistent with the ICAPM when tested over industry returns. Following MSC, the relative risk aversion (RRA) can be estimated by dividing the market beta risk price by the variance of the market excess return (see Eq. (B.11) in Appendix B). The annualized market beta risk price associated with C is 7.31% in Panel A of Table 27 and the annualized standard deviation of the market excess return (proxied by the first-level index of the Datastream Global Equity Indices for Europe, see Sec. 3.3) is 16.41%. Hence, the RRA estimate is 2.71 and thus between one and 10, such that C fulfills ICAPM criterion 1 when its factor loadings are estimated using full-period regressions. Moreover, the signs of the factor beta risk prices associated with size, value, and momentum are the same as the signs of the respective covariance risk prices when the model is estimated over SBM25 displayed in Panel A of Table 18 (p. 107), that is, negative, positive, and positive. Since C meets both ICAPM criteria 3a and 3b when tested over these portfolios, it also meets these criteria in the test with the industry returns and full-period regressions. When C is tested using

[48] Fama and French (2013) augment FF3 with two factors that seek to capture return patterns associated with profitability and investment.

60-month rolling regressions, the RRA estimate is 2.21, so that the model again meets ICAPM criterion 1. As is the case with SM25, the model, however, does not meet ICAPM criterion 3a, since the signs of the (beta) risk prices are inconsistent with the signs of the predictive regressions slopes in Table 14 (p. 99). The model nevertheless fulfills ICAPM criterion 3b when one interprets the indeterminate signs of the predictive regression slopes in Table 16 (p. 102) as positive, positive, and negative. Hence, the results with regard to C's consistency with the ICAPM that are obtained with SBM25 and SM25 as test assets are, all in all, robust to using industry portfolios instead.

Third, it looks as if the theoretical foundations of C are still largely unclear. Why should the five-year cumulative sum on the momentum factor (the variable that is used in the predictive regressions associated with Research Question 3) be a state variable that describes the average investor's set of investment opportunities? This variable has never been proposed in the predictability of returns literature and there is no theoretical model that tells us so.[49] Moreover, Fama and French (2013, p. 5) term the ICAPM justification of their five-factor model, which is an empirical asset pricing model designed similarly to C, "the more ambitious interpretation."[50] They emphasize that their size and value factors do *not* represent state variable-mimicking portfolios. Instead, they are just factors that capture the effects that (up to four, in case of a five-factor model) *unknown* state variables have on expected returns. A critic may argue that it is difficult to convince a decision maker or regulatory authority of a model whose interpretation or theoretical justification is (still) that elusive.

Fourth, like Griffin (2002), Hou et al. (2011), Fama and French (2012), among others, my evidence from the European stock market and industries ignores exchange rate risk, at least in the pre-euro period. As emphasized by Fama and French (2012), this means an implicit assumption of either complete purchasing power parity or that the assets under consideration cannot be used to hedge exchange rate risk. Hence, exchange rate risk potentially biases my results. Finally, one may argue that it is debatable whether it is the right approach to estimate CEs for European industries using pan-European risk factors instead of adopting a more country-specific perspective by separating industries by countries, since this approach assumes that factor loadings and risk premiums do not vary across countries. Griffin (2002, p. 798) examines country-specific and global versions of FF3 and concludes that "cost-of-capital calculations, performance measurement, and risk analysis using Fama and French-style models are best done on a within-country basis."[51] The author does, however, also admit that "it is important to note that our findings do not directly speak to a wider array of models. Better risk proxies may ascribe a more important role to global factors" (p. 798)."

The analysis conducted in this thesis investigates a variety of risk factors in addition to those included in FF3. In addition, I do not construct world factor models as Griffin does for FF3. Instead, I investigate region-specific models, which makes a considerable difference. My

[49] One could argue that the momentum factor is an important part of the average investor's investment opportunity set and its five-year cumulative sum indicates the relative level of its expected return.

[50] The less ambitious interpretation leans on the work of Huberman and Kandel (1987) and suggests that the mean–variance-efficient tangency portfolio that prices all assets consists of the risk-free rate and the model's five traded factors.

[51] I thank two anonymous referees for pointing this out.

approach follows Fama and French (2012), who investigate region-specific versions of the CAPM, FF3, and C versus worldwide versions of these models, terming the region-specific ones local. Furthermore, Bekaert et al. (2009) investigate time trends in country return correlations and, in contrast to other regions, find a significant upward trend for stock return correlations in Europe probably due to the increasing financial and economic integration that is associated with the European Union. The authors' finding speaks in favor of taking a European perspective instead of a country-specific view.

7. Summary and Concluding Remarks

Much empirical evidence suggests that expected asset returns vary through time and that a great amount of variation in expected returns across assets cannot be explained by the CAPM of Sharpe (1964) and Lintner (1965a).[52] Instead, it looks as if the explanation of anomalies such as size, value, and momentum requires asset pricing models that comprise multiple factors. The Intertemporal CAPM (ICAPM) of Merton (1973) provides a theoretical connection between time-varying expected returns and multiple factors in the cross section of expected returns. However, previous empirical evidence from the US stock market, provided by Maio and Santa-Clara (2012, hereafter MSC), suggests that many multifactor models are actually not consistent with the ICAPM, including those that are explicitly justified by their authors as empirical applications of this theory. A large portion of this empirical evidence is based solely on US stocks, whereas evidence from other stock markets and asset classes is relatively sparse. The main motivation of this thesis is to test the out-of-sample (OOS) validity of this evidence using data from the markets of commodity futures and European stocks. Subsequent to a literature review, I formulate five research questions.

Regarding the first research question, I test whether the null hypothesis of unpredictable aggregate returns on commodity futures can be rejected and I attempt to identify variables that show predictive power over these returns. For this purpose, I propose a set of 32 candidate predictors that include stock, bond, macroeconomic, and commodity characteristics and test both their in-sample (IS) and OOS forecasting abilities. The results suggest that many of the candidate predictors have IS predictive power over short- and long-horizon commodity futures returns. Moreover, they indicate that it is possible to forecast returns on commodity futures OOS, especially through the forecast combining methods proposed by Rapach et al. (2010), although the majority of individual forecasts as well as a model selection procedure perform rather poorly in terms of OOS R^2. Hence, these results indicate that the null hypothesis, that is, commodity futures returns are unpredictable, can be rejected. Rather, the results support the alternative hypothesis that expected returns on commodity futures depend on factors such as price levels and past price movements, economic conditions, and investor sentiment and thus vary over time. Overall, the evidence can be interpreted as one more data point that supports a rejection of the null hypothesis that asset returns are *generally* unpredictable.

For the second and third research questions, I conduct a European investigation of eight multifactor asset pricing models—Pástor and Stambaugh's (2003) four-factor model (PS), Campbell and Vuolteenaho's (2004) unrestricted version of the ICAPM (CV), Hahn and Lee's (2006) ICAPM (HL), Petkova's (2006) ICAPM (P), the three-factor model of Koijen et al. (2010, KLVN), Fama and French's (1993) three-factor model (FF3), Carhart's (1997) four-factor model (C), and Fama and French's (1993) five-factor model (FF5)—that were previously tested by MSC using US data in an attempt to assess their OOS validity. For this purpose, I use a large sample of stocks from 16 European countries. These countries account

[52] This chapter is based on Secs. 1 and 6 of Lutzenberger (2014a), Secs. 1 and 7 of Lutzenberger (2014b), as well as Secs. 1 and 5 of Lutzenberger (2014c).

for, on average, 30% of global market capitalization and represent the second largest integrated stock market region, after North America (Fama and French, 2012). In particular, I test their ability to explain the cross section of the average excess returns on 25 portfolios sorted on size and the book-to-market ratio and on 25 portfolios sorted on size and momentum in terms of factor significance, R^2 values, and mean absolute pricing errors. Moreover, I test whether the eight models under investigation meet the restrictions emphasized by MSC, which must be satisfied by a multifactor model to be justifiable as an empirical application of the ICAPM.

The results suggest that most of the multifactor models investigated do a reasonably good job of explaining the testing portfolios' mean excess returns. The results suggest that models with empirical factors, particularly C, have, on average, greater explanatory power within the European stock market than models with economic factors. In particular, HL and KLVN, which include economic factors, seem to be unable to explain the average excess returns of the testing assets, in contrast to MSC's US evidence. Nevertheless, P and CV, which also include economic factors, seem to keep up with the empirical factor models. For example, P performs as well as C in empirically explaining the momentum anomaly when tested over the European sample, whereas, regarding MSC's results, C's good performance in explaining the momentum effect seems to be unbeatable within US data. Yet, all in all, the models' empirical explanatory power seems to be slightly lower for the European stock market.

Moreover, the results do, overall, confirm the US evidence, in that most models do not seem to be justifiable as empirical applications of Merton's theory. Only C seems to be consistent with the ICAPM when tested over both the 25 European portfolios sorted on size and book to market and the 25 European portfolios sorted on size and momentum. The high consistency of C with the ICAPM coincides with the US evidence, although the results are not similar regarding the testing assets and the ICAPM criteria that are satisfied. However, FF3, in contrast to the US evidence, does not seem to be an empirical version of the ICAPM when tested over the European data set. Instead, KLVN seems to be justifiable by the ICAPM when I consider investment opportunities that are driven by changes in market volatility (although only when tested over the size/book-to-market portfolios), which is not the case in the US study. Moreover, HL seems to be consistent with the ICAPM when tested over the size/momentum portfolios and when I consider investment opportunities that are driven by changes in expected market returns, again in contrast with the US results.

These results are quite robust to a set of methodological changes. To be specific, my assessment of the multifactor models' explanatory powers for the cross section of average excess returns seems to be relatively robust, except when an intercept is included in the pricing equations, when bond risk premiums are added to the test assets, or when the models are estimated in expected return–beta form via generalized least squares. Moreover, especially the assessment that C can be justified by the ICAPM appears to be relatively insensitive with respect to adjustments in the underlying empirical methodology.

Finally, the fourth and fifth research questions examine cost of equity capital (CE) estimates for European industries that are obtained from the CAPM, as well as from the eight multifactor models that are investigated earlier in this thesis. In particular, I build on the studies of Fama and French (1997) and Gregory and Michou (2009), who assess CE estimates for US and UK industries that are obtained from the CAPM, FF3, as well as C, in two ways. First, I extend these studies to 35 European industries to assess their robustness using international data. Second, I extend them with the six remaining multifactor models that I examine earlier. These additional models include mainly economically motivated factors, such as liquidity-related variables or innovations in economic state variables, rather than purely empirical factors. Moreover, their additional factors are all non-traded. Risk premiums for non-traded factors cannot be estimated by simply using historical averages. Instead, these non-traded factors require a different methodology for CE estimation than the method that is usually applied for traded factors. I propose such a methodology, which is based on Fama and MacBeth's (1973) cross-sectional regressions.

Rather unsurprisingly, the results suggest that CE estimates obtained from the CAPM, FF3, and C are as inaccurate for European industries as they are for US and UK industries. First, I obtain imprecise estimates of factor risk loadings because the true loadings vary considerably through time. Furthermore, forecast results indicate that full-period risk loading estimates should be preferred to rolling regression estimates to discount the typical industry's cash flows at all horizons. Second, estimates of factor risk premiums are imprecise as well. For example, if one estimates the market equity premium by its sample mean, the true premium might be anywhere between -2.3% and 11.9% according to the rule of thumb of plus or minus two standard errors. Unfortunately, the six models that I additionally examine seem to do more harm than good because their CE estimates seem to be even more imprecise than those obtained from the CAPM, FF3, and C. Many risk factors in these models are non-traded. The variability through time of the true loadings on these non-traded risk factors is substantially higher and the errors from predicting loadings 60 months ahead using past returns are much greater for most of these factors. Moreover, the premiums that I obtain from Fama–MacBeth (1973) regressions for these non-traded factors are no more precise than the premiums for the traded size, value, and momentum factors of Fama and French (1993) and Carhart (1997). Consequently, these models are not only inconsistent with the ICAPM in the US according to MSC and in Europe, with the exception of KLVN and HL, according to my results, but they also seem to be less qualified to estimate precise CEs. Finally, the results show that the use of various (multi-)factor asset pricing models, methods for the estimation of risk loadings, and procedures to estimate risk premiums result in highly different CE estimates for individual industries. Some of these deviations are even distressing and are probably even greater for individual firms. Hence, the choice of model appears to be important.

These results suggest that if European decision makers and regulatory authorities insist upon employing a (multi-)factor model, they should use C for practical purposes such as CE estimation for several reasons. First, the model shows the greatest explanatory power for the cross section of expected European stock returns. Second, it is most likely consistent with the ICAPM in Europe. Third, it provides CEs that are more precise than those obtained from multifactor models that comprise primarily economic factors. Nevertheless, the results of Fama and French (1997), Gregory and Michou (2009), and this thesis indicate that the application of (multi-)factor models for CE estimation is highly questionable. In his Presidential Address to the American Finance Association, Cochrane (2011) predicts that the

CAPM's empirical failures will deeply change the applications of asset pricing models within corporate finance, accounting, and regulatory authorities that still commonly use the CAPM. However, Cochrane also emphasizes that he does not see a solution to practical problems such as project valuation in multifactor models. Instead, the author advocates the use of simple "comparables" such as average returns on similar securities. My results support this conclusion.

A task for future research is to connect the predictability patterns that I find in the commodity futures market to established theories, for instance, to the theory of storage of Kaldor (1939) and others, and to refine these theories to account for observed empirical patterns where required. Moreover, the finding that investor sentiment seems to be relatively successful in predicting commodity futures returns is quite interesting. This result brings the field of behavioral finance into play against theories of efficient markets and rational investors. I do, however, present this factor with just one out of many different possible proxies (for an overview, see Baker and Wurgler, 2007). Hence, predictability tests of commodity futures returns with further sentiment proxies and other potential predictive variables from behavioral finance would be interesting. Furthermore, I suggest providing empirical tests of the remaining multifactor models that are presented in the literature review of this thesis with regard to their ability to explain average returns on European stocks, their consistency with the ICAPM, as well as their applicability in estimating CEs. In addition, these remaining multifactor models should also be tested on their consistency with the ICAPM in the US stock market. Beyond that, it might make sense to improve the conducted cross-sectional asset pricing tests by implementing more of the suggestions of Lewellen et al. (2010) or Daniel and Titman (2012). For instance, one could expand the set of test assets by portfolios sorted by, for instance, factor loadings. Additionally, it would be interesting to further expand the study to other stock market regions, such as Japan and the Asia Pacific. Last but not least, we have to find a reasonable model or methodology to obtain accurate CE estimates that take the empirical observations of time-varying expected returns and deviations from the CAPM in the cross section of returns adequately into account.

To conclude this thesis, I would like to quote one of the 2013 Nobel Prize recipients, Fama (1991, p. 1610), whose statement appears to be—after more than 20 years of further research—(unfortunately) still up to date:

> In the end, I think we can hope for a coherent story that (1) relates the cross-section properties of expected returns to the variation of expected returns through time, and (2) relates the behavior of expected returns to the real economy in a rather detailed way. Or we can hope to convince ourselves that no such story is possible.

Appendix A. GMM Formulas

The GMM methodology follows Cochrane (2005, Sec. 13.2) and MSC.[53] The GMM system consists of $N + K + 1$ moment conditions. The first N moments are the pricing errors of the N test assets. The last K + 1 moment conditions are used to estimate the factor means, where K is the number of factors in addition to the market factor. The most important formulas are outlined in the following. In general, a factor model can be formulated as

$$\begin{gathered} E\left[R^e\left(1-\left(f-E(f)\right)'b\right)\right]=0, \\ E(R^e)=Cov(R^e f')b, \end{gathered} \tag{A.1}$$

where R^e is an Nx1 vector of excess returns, f is a (K + 1)x1 vector of factors, and b is a (K + 1)x1 vector of factor risk prices. Following MSC, the weighting matrix associated with the GMM system is formulated as

$$W=\begin{bmatrix} W^* & 0 \\ 0 & I_{K+1} \end{bmatrix}, \tag{A.2}$$

where W^* is an NxN weighting matrix, 0 represents a conformable matrix of zeros, and I_{K+1} denotes a (K + 1)-dimensional identity matrix. In this formulation, W^* is the weighting matrix associated with the first N moment conditions, which correspond to the N pricing errors of the N testing assets, while I_{K+1} represents the weighting matrix for the last K + 1 orthogonality conditions, which identify the means of the K + 1 factors that each model comprises. In the first-stage GMM, which is conceptually equivalent to an OLS cross-sectional regression of expected excess returns on the covariances between returns and factors, W^* is the identity matrix, $W^* = I_N$. The first-stage GMM estimates of b and $E(f)$ are then

$$\begin{gathered} \hat{b}=[C'C]^{-1}C'E_T(R^e), \\ E(f)=E_T(f), \end{gathered} \tag{A.3}$$

where $E_T(\cdot)$ denotes the sample mean and

$$C \equiv E\left(R^e \tilde{f}'\right) \tag{A.4}$$

denotes the covariance matrix of returns and factors, with $\tilde{f} = f - E(f)$. The variance–covariance matrix of the estimated parameters is

$$Var\left[\hat{b},E(f)\right]=\frac{1}{T}(d'Wd)^{-1}d'W\hat{S}Wd(d'Wd)^{-1}, \tag{A.5}$$

where d denotes the derivative of moments with respect to the parameters b and $E(f)$:

$$d=\begin{bmatrix} -C & E(R^e)b' \\ 0 & -I_{K+1} \end{bmatrix}. \tag{A.6}$$

[53] This chapter is based on Sec. 1 of the Internet Appendix of Lutzenberger (2014b).

In this formulation, $\hat{S}$ is an estimator for the spectral density matrix, S, which is derived under the heteroskedasticity-robust or White (1980) standard errors, that is, under the null hypothesis that pricing errors and factors are not forecastable from past information:

$$\hat{S} = \begin{bmatrix} E(u_t u_t') & E(u_t \tilde{f}_t') \\ E(\tilde{f}_t u_t') & E(\tilde{f}_t \tilde{f}_t') \end{bmatrix}, \tag{A.7}$$

where u_t denotes the pricing errors,

$$u_t \equiv R_t^e (1 - \tilde{f}_t' b). \tag{A.8}$$

The variance–covariance matrix of the estimated factor risk prices, $Var(\hat{b})$, is the top left $(K+1)$x$(K+1)$ element of $Var[\hat{b}, E(f)]$. We can use $Var(\hat{b})$ to test the statistical significance of each risk price; that is, we can test whether the risk price of the ith factor is equal to zero (Cochrane, 2005, p. 192):

$$\frac{\hat{b}_i}{\sqrt{Var(\hat{b})_{ii}}} \sim N(0,1). \tag{A.9}$$

Appendix B. Sensitivity Analysis

I assess the robustness of the results associated with Research Questions 2 and 3 by making several methodological adjustments.[54] First, I make methodological changes with regard to the predictive regressions (Secs. B.1 and B.2). Second, I change the cross-sectional methodology (Secs. B.3 to B.9). For a better presentation, I display the tables that contain the results associated with all methodological adjustments at the end of Sec. B.9.

B.1 Alternative Proxies for the State Variables Associated with SMB and HML

I employ different proxies for the state variables associated with SMB and HML in the predictive regressions, $CSMB$ and $CHML$, following Sec. 5.6 of MSC. The alternative proxies are constructed similarly to CL and $CUMD$. For instance, $CSMB$ is constructed as

$$CSMB_t = \sum_{s=t-59}^{t} SMB_s. \tag{B.1}$$

This methodological change is only relevant for FF3, C, PS, and FF5.

The results for the return-predicting regressions are displayed in Table 30 (p. 158). I compare the overall assessment of the signs of the predictive slope coefficients (Panel D) to the benchmark case with SMB^* and HML^* (Panel D of Table 14, p. 99). In contrast to the benchmark case, $CSMB$ forecasts aggregate returns significantly negatively in FF3. Moreover, the signs of the slopes of $CHML$ and $CUMD$ are both indeterminate in C. The slope of $CUMD$ is assessed as indeterminate, since the sign of the slope at $q = 36$, which is not tabulated here, is negative (and non-significant). Furthermore, the sign of the slope of $CHML$ is indeterminate in PS. Finally, in FF5, the signs of the slopes of $CSMB$ and $CHML$ are both indeterminate and the slope of DEF is now significantly positive. In the second step, I compare the overall assessment of the signs of the predictive slopes to the signs of the corresponding factor risk premiums (Table 18, p. 107). Regarding the SBM25 portfolios (Panel A), in contrast to the original assessment, the risk prices of FF3 satisfy the sign restriction implied by the ICAPM and hence FF3 satisfies ICAPM criterion 3a. C still satisfies criterion 3a, but only when I assume the indeterminate signs of the slopes of $CHML$ and $CUMD$ are positive. Moreover, while PS still does not satisfy criterion 3a, the criterion is now met by FF5 when I interpret the indeterminate signs of the slopes of $CSMB$ and $CHML$ as positive. With regard to the SM25 portfolios, observe that criterion 3a remains unmet by all four relevant models.

The outcomes for the volatility-predicting regressions are shown in Table 31 (p. 160). Note that, in contrast to the benchmark case with SMB^* and HML^* (Panel D of Table 16, p. 102), the slope of $CSMB$ is significantly positive in FF3. Moreover, in C, the slope of $CSMB$ is significantly positive too, while the slope of $CHML$ is significantly negative. Finally, in FF5, the signs of the slope of both $CSMB$ and DEF are now indeterminate, while the slope of

[54] This chapter is based on Sec. 2 of the Internet Appendix to Lutzenberger (2014b).

$TERM$ is now significantly negative. Regarding the factor risk premiums estimated over SBM25 (Panel A of Table 18, p. 107), note that, in contrast to the original assessment, FF3 now satisfies ICAPM criterion 3b. Moreover, observe that C still satisfies criterion 3b when one interprets the indeterminate sign of the slope of $CUMD$ as negative. Furthermore, while PS still does not meet criterion 3b, FF5 now satisfies the sign restrictions of criterion 3b when one assumes the indeterminate signs of the slopes of $CSMB$ and DEF to be negative. With regard to the factor risk premiums estimated over SM25 (Panel B of Table 18, p. 107), observe that, in contrast to the original assessment, C no longer meets criterion 3b.

B.2 Alternative Volatility Measures

I use three alternative volatility measures in the volatility-predicting regressions (MSC, Sec. 3.1 of the Internet Appendix). First, I employ the volatility measure proposed by Beeler and Campbell (2012). For this purpose, I conduct a one-month predictive regression on the relevant state variables of each multifactor model in the first step:

$$r_{t+1} = \kappa_0 + \kappa_1' z_t + u_{t+1}. \tag{B.2}$$

In the second step, I perform a regression for each multifactor model that contains q-period cumulative squared residuals:

$$\frac{\sum_{j=1}^{q} u_{t+j}^2}{q} = a_q + b_q' z_t + \xi_{t,t+q}. \tag{B.3}$$

Tables 32 (p. 161) and 33 (p. 162) show the results and Panel D displays the overall assessments based on all seven forecasting horizons.[55] With regard to the overall assessment, observe that, in contrast to the original assessment with $SVAR$ (Panel D of Table 15, p. 100, and Table 16, p. 102), the sign of the slope on DEF is indeterminate in HL. The same applies to the signs of the slopes of RF (P), CP (KLVN), and DEF (FF5). Moreover, the slope of PE is significantly positive in CV and that of SMB^* is significantly negative in C. Comparing these results to the signs of the factor risk premiums for the ICAPM specifications estimated over the SBM25 portfolios (Panel A of Table 17, p. 105), one can see that, in contrast to the benchmark case, HL satisfies ICAPM criterion 3b when one assumes the indeterminate sign of the slope of DEF to be negative. Moreover, KLVN satisfies criterion 3b only when one interprets the indeterminate sign of the slope of CP as positive. With regard to the SM25 portfolios (Panel B of Table 17, p. 105), note that P only meets criterion 3b when the indeterminate sign of the slope of RF is assumed to be negative. Furthermore, KLVN satisfies criterion 3b when one interprets the indeterminate sign of the slope of CP as negative. Regarding the premiums for the empirical risk factors estimated over SBM25 (Panel A of Table 18, p. 107), one observes that, in contrast to the original assessment, C does not fulfill

[55] In Panel D of Table 33 (p. 162), the assessment +/- of the sign of the slope of SMB^* within PS is not obvious from considering only the results for the one-, 12-, and 60-month horizons, which are shown in the tables. The assessment comes from the fact that SMB^* forecasts volatility at horizons of one month, three months, and 12 months to be significantly negative, while it predicts volatility at the 36-month horizon (not tabulated here) to be significantly positive.

ICAPM criterion 3b. When one looks at SM25 (Panel B of Table 18, p. 107), one can see that there is no change in the original assessment with regard to ICAPM criterion 3b.

Second, I use the volatility measure of Bansal et al. (2005), for which I specify a first-order autoregressive AR(1) process for the log market return in the first step:

$$r_{t+1} = \phi_0 + \phi_1 r_t + u_{t+1}. \tag{B.4}$$

In the second step, I conduct for each multifactor model the following volatility-predicting regression:

$$ln\left(\sum_{j=1}^{q}\left|u_{t+j}\right|\right) = a_q + b_q' z_t + \xi_{t,t+q}. \tag{B.5}$$

Tables 34 (p. 163) and 35 (p. 164) show the results. With regard to the overall assessment (Panel D), note that, in contrast to the original assessment with $SVAR$ (Panel D of Table 15, p. 100, and Table 16, p. 102), the sign of the slope of DEF is indeterminate in HL, which also applies to the signs of the slopes of $TERM$ (P) and DEF (FF5). Moreover, the slopes of RF (P) and PE (CV) are significantly positive, while the slopes of CP (KLVN), SMB^* (C and PS), HML^* (C), and $CUMD$ (C) are now significantly negative.[56] Comparing these results to the factor risk premiums of the ICAPM specifications (SBM25), one can see that HL now satisfies ICAPM criterion 3b when one assumes the indeterminate sign of the slope of DEF to be negative. However, KLVN no longer meets criterion 3b. Regarding SM25, I note that, on the one hand, P no longer satisfies criterion 3b. On the other hand, this criterion is now met by KLVN. Looking at the premiums for the empirical risk factors, one can see that C no longer meets criterion 3b over SBM25, while the criterion is met by the model when it is estimated over SM25.

Third, I employ the conditional second moment of aggregate returns, that is, the monthly squared continuously compounded market return, as an alternative volatility measure and conduct the following predictive regressions:

$$r_{t,t+q}^2 = a_q + b_q' z_t + v_{t,t+q}. \tag{B.6}$$

The results are displayed in Tables 36 (p. 165) and 37 (p. 166). Panel D shows the overall assessment.[57] Comparing the overall assessment to the original assessment with $SVAR$, one

[56] In Table 34 (p. 163), Panel D, the assessment + of the predictive slope on RF within the P model is due to the fact that RF predicts volatility at the 36- and 48-month horizons (not tabulated here) significantly positively.

[57] In Table 36 (p. 165), Panel D, the plus sign with regard to the slope of $TERM$ (HL) comes from the fact that $TERM$ forecasts volatility significantly positively at horizons of 24 months, 36 months, and 48 months (not tabulated here). The same applies to the plus sign for $TERM$ within CV. The +/- sign for DEF (P) is a result of the fact that DEF predicts volatility at the one- and three-month horizons significantly positively, while it forecasts volatility at the 36-month horizon (not tabulated here) significantly negatively. In Table 37 (p. 166), Panel D, the +/- sign for CL (PS) is a consequence of the fact that CL predicts volatility at the one-month horizon significantly negatively, while it forecasts volatility at the 24- and 36-month horizons (not tabulated here) significantly positively. In the same panel, the plus sign with regard to $TERM$ (FF5) is due to the fact that $TERM$ is a significantly positive predictor of volatility at the 24-, 36-, and 48-month horizons (not tabulated here) and the +/- sign regarding DEF (FF5) results from the fact that DEF forecasts volatility at the

can see that the slopes of $TERM$ (P), DY (P), RF (P), VS (CV), HML^* (FF3, C, PS, and FF5), and $CUMD$ (C) are now significantly positive, while the slopes of PE (CV), CP (KLVN), and SMB^* (C and PS) are now significantly negative. Moreover, the signs of the slopes of CL (PS) and DEF (FF5) are, in contrast to the benchmark case, indeterminate.

I then compare these results to the factor risk premiums for the ICAPM specifications (Table 17, p. 105). With regard to SBM25, note that CV and KLVN no longer satisfy ICAPM criterion 3b. Looking at the SM25 risk premiums, one can see that KLVN now meets criterion 3b, while the criterion is no longer met by P and CV. Regarding the premiums for the empirical risk factors (estimated over SBM25), observe that C no longer meets criterion 3b. Regarding the SM25 results, I note that FF3 now satisfies criterion 3b, while it is no longer met by C. Moreover, PS now meets criterion 3b when one assumes the indeterminate sign of the slope of CL to be positive and FF5 now satisfies the criterion when one interprets the indeterminate sign of the slope of DEF as negative.

B.3 Including an Intercept in the Pricing Equations

I include an intercept in the pricing equation of each multifactor model and then re-estimate all the models (MSC, Sec. 5.1). First, I regard the results for the ICAPM specifications, which are shown in Table 38 (p. 167). Note that the intercept is economically and statistically significant within all models and over both SBM25 and SM25 (varying between 1% and 2% per month). Moreover, all models show higher R^2_{OLS} and lower MAE values than in the original tests. Especially with regard to SM25, I observe that the R^2_{OLS} values of most models, that is, of the CAPM, HL, CV, and KLVN, increase substantially in comparison to the benchmark values (Panel B of Table 17, p. 105). These findings indicate that the models are not specified correctly. Furthermore, in contrast to the benchmark case, HL and KLVN show RRA estimates that are negative (over both SBM25 and SM25), so that these models no longer satisfy ICAPM criterion 1. With regard to the factor risk prices, I observe that γ_{CP} within KLVN (SBM25) and γ_{DEF} within P (SM25) are now positive. Consequently, KLVN no longer satisfies ICAPM criterion 3b when tested over SBM25, while P no longer meets criterion 3a when tested over SM25. However, P still satisfies criterion 3b over SM25 when the indeterminate sign of the slope of DEF in the volatility-predicting regressions (Panel D of Table 15, p. 100) is assumed to be negative.

Second, I look at the outcomes for the models with empirical risk factors, which are displayed in Table 39 (p. 168). The intercept is economically and statistically significant within all models, except in C estimated over SM25 (varying between 1% and 2% per month). Moreover, the R^2_{OLS} values of FF3 and PS, estimated over SM25, increase substantially in comparison to the original estimation, indicating that the models are not correctly specified. Observe that the RRA estimates of FF3, PS (over both SBM25 and SM25) and FF5 (over SBM25) are now negative, in contrast to the benchmark case. Hence, ICAPM criterion 1 is no longer met in these tests. However, FF5 now satisfies criterion 1 when the model is tested

one- and three-month horizons significantly positively, while it predicts volatility at the 36-month horizon (not tabulated here) significantly negatively.

over SM25, with an RRA estimate of 1.89 (but non-significant). Moreover, note that γ_{SMB} is, in contrast to the original estimation, negative within FF3 (SM25), C (SM25), PS (SBM25 and SM25), and FF5 (SBM25). Hence, the estimation of this parameter seems to be very sensitive to the inclusion of an intercept in the pricing equations. In addition, γ_L is now positive when PS is estimated over SM25. The consequences of these sign changes are that C now meets ICAPM criterion 3a (SM25). Moreover, C only satisfies criterion 3b over SM25 when the indeterminate sign of the slope on SMB^* is assumed to be positive.

B.4 Estimating the Asset Pricing Models with Second-Stage GMM

All the multifactor models are re-estimated with second-stage GMM (MSC, Sec. 5.2 and Appendix A, and Cochrane, 2005, Sec. 13.2). The second-stage GMM estimate is equivalent to a GLS cross-sectional regression of expected excess returns on the covariances between returns and factors, but the GMM standard errors are corrected for the fact that the factor means are estimated. In the second-stage GMM, the weighting matrix associated with the first N moments is the inverse of the first NxN element of the spectral density matrix, $W^* = S_N^{-1}$. The second-stage GMM estimate of b is

$$\hat{b} = [C' S_N^{-1} C]^{-1} C' S_N^{-1} E_T(R^e). \tag{B.7}$$

The remaining formulas are the same as in Appendix A. Note that the estimate from the GLS cross-sectional regression does not correspond to the fully efficient GMM estimate. In case of the fully efficient GMM estimation, the weighting matrix would be represented by the inverse of the full spectral density matrix, S^{-1} (Maio, 2013). I follow MSC and compute the weighted least squares (WLS) coefficient of determination, which is

$$R_{WLS}^2 = 1 - \frac{\bar{\hat{\alpha}}' \widehat{S^*}_N^{-1} \bar{\hat{\alpha}}}{\bar{R}' \widehat{S^*}_N^{-1} \bar{R}}, \tag{B.8}$$

where $\bar{\hat{\alpha}}$ denotes the vector of demeaned pricing errors, $\widehat{S^*}_N$ includes the diagonal elements of S_N, and $\bar{R}$ is the vector of demeaned (average) excess returns.

I first consider the second-stage GMM estimation results for the ICAPM specifications, shown in Table 40 (p. 169). Note that the R_{WLS}^2 estimates hardly differ from the R_{OLS}^2 estimates in the benchmark tests (Table 17, p. 105). Moreover, observe that P, in contrast to the benchmark case, shows an RRA estimate between one and 10 over SBM25, so that the model meets ICAPM criterion 1. However, HL now shows an RRA estimate that is slightly above 10 when the model is estimated over SM25, so it no longer meets criterion 1. Furthermore, note that, in contrast to the first-stage GMM estimation, γ_{TERM} is positive within all models and over both sets of test assets. In addition, γ_{VS} is now negative when CV is tested over SM25. Consequently, HL now satisfies criterion 3a over SBM25 when the indeterminate sign of the slope of DEF within the return-predicting regressions (Panel D of Table 13, p. 97) is assumed to be positive. In addition, KLVN now meets criterion 3a over SBM25. However, CV no longer meets criterion 3b over either SBM25 or SM25.

Second, I look at the results for the models with empirical risk factors, reported in Table 41 (p. 170). Note that the explanatory ratios again hardly differ from those of the benchmark tests. Moreover, note that the RRA estimates of C, PS, and FF5 are, in contrast to the first-stage GMM estimation results, now slightly above 10 when these models are tested over SBM25, so that criterion 1 is no longer met in these tests. Furthermore, observe that the signs of the risk prices differ only in two estimations, that is, γ_{SMB} is negative within C and γ_{HML} is positive within FF5, both when estimated over SM25. The consequence of these sign changes is that C now satisfies criterion 3a over SM25. Moreover, the model meets criterion 3b over SM25 only when the indeterminate sign of the slope of SMB^* in the volatility-predicting regression (Panel D of Table 16, p. 102) is assumed to be positive.

B.5 Adding Excess Bond Returns to the Test Assets

I add the excess returns on seven artificial German government bonds with residual maturities of one year, two years, five years, seven years, 10 years, 15 years, and 20 years to both sets of testing assets and re-estimate each multifactor model (MSC, Sec. 5.3). I thus have a total of 33 testing assets (25 stock portfolios, the equity premium, and seven bonds) in each test. The bond returns are obtained using the term structure of interest rates published by Deutsche Bundesbank (series WZ9801, WZ9802, WZ9803, WZ9804, WZ9805, and WZ9806). MSC employ Treasury bonds with maturities of one year, two years, five years, seven years, 10 years, 20 years, and 30 years. However, due to the bond issuance policies of the German government, the time series for German government bonds with maturities greater than 20 years are not long enough for my sample period, so I use the 15-year instead of the 30-year artificial bond.

Table 42 (p. 171) shows the results for the ICAPM specifications. Note that although the pricing of both stock and bond risk premiums at the same time should be more demanding for the models, the explanatory ratios do not change that much in comparison to the benchmark tests (Table 17, p. 105). Moreover, observe that, in contrast to the benchmark case, γ_{TERM} is negative within HL and P (SM25), γ_{DEF} is negative within HL (SBM25), and γ_{CP} is negative within KLVN (SM25). The consequences of these sign changes are that HL now satisfies ICAPM criterion 3b over SBM25, while criterion 3a is no longer met by the model over SM25. Moreover, P no longer meets either criterion 3a or 3b over SM25. However, KLVN now satisfies criterion 3b over SM25.

Table 43 (p. 172) displays the results for the models with empirical risk factors. Note that the explanatory ratios of most models are quite robust with respect to the benchmark tests (Table 18, p. 107). The greatest changes are shown by the explanatory ratios of FF3 (over SBM25) and C (over both SBM25 and SM25), which both seem to have noticeably more difficulties in simultaneously explaining both stock and bond risk premiums. Observe that, in contrast to the original estimation without excess bond returns, the RRA estimate of PS is slightly above 10 (SBM25 and SM25). Hence, ICAPM criterion 1 is no longer fulfilled by the model in these tests. Moreover, γ_{SMB} within C is now negative (SM25), γ_{HML} within FF5 is positive (SM25), and γ_{TERM} within FF5 is negative (SBM25). The result of these different risk price signs is that C now satisfies criterion 3a over SM25. Moreover, the model meets criterion 3b over

SM25 when the indeterminate sign of the slope on SMB^* in the volatility-predicting regressions is assumed to be positive.

B.6 Excluding the Market Excess Return from the Test Assets

I exclude the market excess return from both sets of testing assets and re-estimate each multifactor model (MSC, Sec. 2.2 of the Internet Appendix). For one thing, this adjustment shows how large the impact is of forcing the models to price one of the factors on the results. For another thing, most cross-sectional asset pricing tests do not include the equity premium in the set of testing assets. The outcomes displayed in Table 44 (p. 173) show that the explanatory ratios of the ICAPM specifications are quite similar to those in the benchmarks tests. Consequently, the models do not seem to have more or fewer difficulties when the equity premium is not part of the set of testing assets. Moreover, in contrast to the benchmark tests with the equity premium, the RRA coefficient of KLVN is slightly below one when I test the model over SM25. Hence, the model no longer meets ICAPM criterion 1. Furthermore, observe that, in contrast to the original tests, γ_{TERM} is positive within KLVN (SBM25) and negative within P (SM25). Moreover, γ_{DEF} is negative within HL (SBM25) and γ_{PE} is positive within CV (SM25). Consequently, HL now fulfills criterion 3b over SBM25. Moreover, KLVN now meets criterion 3a over SBM25, while the model no longer satisfies criterion 3b over SBM25. Furthermore, P no longer meets criterion 3a or 3b over SM25. Table 45 (p. 174) presents the results for the models with empirical risk factors. The explanatory ratios do not, again, differ much from the benchmark tests with the equity premium, although it seems that most models have slightly greater explanatory power when the equity premium is excluded. With regard to the factor risk prices, the only change that occurs is that γ_{TERM} is now negative within FF5 (SBM25). However, the model still does not meet the ICAPM criteria 3a and 3b.

B.7 Estimating the Asset Pricing Models in Expected Return–Beta Representation

All multifactor models are re-estimated in expected return–beta form (MSC, Sec. 5.4). I employ the time-series/cross-sectional regression approach proposed by Brennan et al. (2004), Cochrane (2005, Sec. 12.2), and others, that is, factor loadings (betas) are estimated by time-series regressions of excess returns on factors for each test asset in the first step. For instance, in the case of HL, I conduct the following regressions:

$$R_{i,t+1} - R_{f,t+1} = \delta_i + \beta_{im} RM_{t+1} + \beta_{iTERM} \Delta TERM_{t+1} + \beta_{iDEF} \Delta DEF_{t+1} + \varepsilon_{i,t+1}, \tag{B.9}$$

where the slope coefficients β_i represent the factor loadings. The second step comprises a cross-sectional regression of the test assets' average excess returns on the estimated factor loadings:

$$\overline{R_i - R_f} = \lambda_M \beta_{im} + \lambda_{TERM} \beta_{iTERM} + \lambda_{DEF} \beta_{iDEF} + \alpha_i, \tag{B.10}$$

where the slope coefficients λ are the factor (beta) risk prices. The cross-sectional regressions are first conducted using the OLS methodology and the t-statistics of the factor (beta) risk prices are computed with Shanken's (1992) approach to account for the estimation error in the factor loadings. The implied RRA estimate is obtained by dividing the market (beta) risk price by the variance of the market excess return:

$$\gamma = \frac{\lambda_M}{Var(RM_{t+1})}. \tag{B.11}$$

Table 46 (p. 175) shows the outcomes for the ICAPM specifications. Observe that the explanatory ratios are the same as in the benchmark (first-stage GMM) estimation (Table 17, p. 105). Moreover, note that, in contrast to the benchmark tests, all models now fulfill ICAPM criterion 1, showing an RRA estimate that is between one and 10. Furthermore, observe that λ_{TERM} is negative within P and positive within KLVN (SM25), which is in contrast to the risk prices obtained in the benchmark tests. The same applies to the positive sign of λ_{DEF} within P and λ_{PE} within CV (SM25). The consequences of these sign changes are that P no longer fulfills ICAPM criteria 3a and 3b when the model is tested over SM25. Moreover, CV only meets criterion 3b when one assumes the indeterminate sign of the slope of PE in the volatility-predicting regressions to be negative (SM25). Table 47 (p. 176) displays the results for the models with empirical risk factors. Observe that each model's RRA estimate is between one and 10, so that ICAPM criterion 1 is met by each model and over both sets of testing assets. Moreover, note that, in contrast to the first-stage GMM estimation results, λ_{SMB} is negative within PS and FF5 (SBM25), as well as within C (SM25). For this reason, C meets criterion 3a over SM25. Moreover, the model meets criterion 3b over SM25 only when the indeterminate sign of the slope on SMB^* in the volatility-predicting regressions is assumed to be positive.

The cross-sectional regressions are repeated using GLS. This method weights the observations according to the inverse of the full covariance matrix of the residuals from the time-series regressions, $\hat{\Sigma}$. The explanatory power of the models is now measured by the GLS coefficient of determination, which is similar to the WLS coefficient of determination, except that the demeaned pricing errors and demeaned average excess returns are weighted by the inverse of $\hat{\Sigma}$:

$$R^2_{GLS} = 1 - \frac{\bar{\hat{\alpha}}'\hat{\Sigma}^{-1}\bar{\hat{\alpha}}}{\bar{R}'\hat{\Sigma}^{-1}\bar{R}}, \tag{B.12}$$

where $\bar{\hat{\alpha}}$ denotes the vector of demeaned pricing errors and $\bar{R}$ is the vector of demeaned (average) excess returns. The results for the ICAPM specifications are displayed in Table 48 (p. 177). Note that the explanatory power of HL, CV (SBM25 and SM25), and P (SM25) improve considerably according to the GLS coefficient of determination. The GLS regression assigns a zero pricing error for each test asset that is also incorporated into the model as a factor. Such a test asset and factor, respectively are the market excess return in all the models I test. Observe that, for this reason, the market (beta) risk price is numerically equal to the mean excess market return in each test, that is, 1.12% per month in my sample, and each model shows the same RRA estimate (3.89). Consequently, ICAPM criterion 1 is fulfilled by each model. Furthermore, note that, in contrast to the benchmark first-stage GMM estimations, λ_{TERM} is positive within HL (SBM25), CV (SM25), and KLVN (SBM25 and

SM25). Moreover, λ_{DEF} is positive and λ_{RF} is negative within P (SM25), λ_{VS} is negative within CV (SM25), and λ_{CP} is positive (SBM25) and negative (SM25) within KLVN. The consequences of these sign changes are that HL now fulfills ICAPM criterion 3a (SBM25) when the indeterminate sign of the slope of DEF in the return-predicting regressions (Panel D of Table 13, p. 97) is assumed to be positive. Moreover, P no longer meets criterion 3a or 3b over SM25 and CV no longer satisfies criterion 3b over SM25. Finally, KLVN no longer satisfies criterion 3b over SBM25. Instead, the model now meets criterion 3a over SM25.

Table 49 (p. 178) displays the results for the models with empirical risk factors. Observe that the models' explanatory power, now measured by R^2_{GLS}, changes noticeably with respect to the benchmark first-stage GMM tests. While the explanatory ratio of C increases over SBM25 but decreases over SM25, the ratios associated with the other three models decrease over SBM25 but increase regarding SM25. Moreover, note that all models now satisfy ICAPM criterion 1. Furthermore, one can see that, in contrast to the benchmark first-stage GMM estimations, λ_{SMB} is negative in all tests. Moreover, λ_{HML} is negative within C and positive within FF5 (SM25). Furthermore, λ_L is positive within PS and λ_{TERM} is positive within FF5 (SM25). The sole consequence of these sign changes is that C only meets criterion 3b over SM25 when the indeterminate signs of the slopes on SMB^* and HML^* in the volatility-predicting regressions are assumed to be positive.

B.8 Estimating the Asset Pricing Models with Orthogonal Factors

I re-estimate each multifactor model by first orthogonalizing each risk factor with respect to the market excess return, following Petkova (2006) and others (MSC, Sec. 2.1 of the Internet Appendix). For this purpose, I conduct the following regression for each factor:

$$f_{t+1} = \psi_0 + \psi_1 RM_{t+1} + u_{t+1}. \tag{B.13}$$

I then compute the new factor as $\hat{f}_{t+1} = \hat{\psi}_0 + \hat{u}_{t+1}$. The new (orthogonalized) factors only add information to the models that is independent of the information associated with the equity premium. The outcomes are displayed in Tables 50 (p. 179) and 51 (p. 180). Note that the explanatory ratios have (by definition) exactly the same values as in the benchmark tests with non-orthogonalized factors. Moreover, observe that all models now show market (covariance) risk prices between one and 10 so that all the models satisfy ICAPM criterion 1. All the other parameters have values that are similar to those from the benchmark tests.

B.9 Different Proxies for the State Variable Innovations

Finally, I employ two alternative methods to compute the state variable innovations of the ICAPM specifications. First, I proxy for the innovations in the state variables using the residuals from an AR(1) process for each state variable (MSC, Sec. 2.4 of the Internet Appendix),

$$z_{t+1} = \phi_0 + \phi_1 z_t + \varepsilon_{t+1}, \tag{B.14}$$

and $\hat{\varepsilon}_{t+1}$ is then used as the estimate for the state variable innovation. Table 52 (p. 181) shows the results for the ICAPM specifications. Observe that the explanatory ratios change only slightly in comparison to the benchmark tests with first differences as proxies for the state variable innovations (Table 17, p. 105). Moreover, note that the RRA estimate of KLVN (SM25) is, in contrast to the benchmark test, slightly below one and, hence, the model does not fulfill ICAPM criterion 1 when tested over SM25. Moreover, regarding the risk price estimates, one can see that γ_{TERM} is now positive within KLVN (SBM25), γ_{DEF} is negative within HL (SBM25), and γ_{PE} is positive within CV (SM25). Consequently, HL now satisfies ICAPM criterion 3b when tested over SBM25 and KLVN meets criterion 3a instead of criterion 3b when estimated over SBM25. Moreover, CV only meets criterion 3b over SM25 when the sign of the indeterminate slope of PE in the volatility-predicting regressions (Panel D of Table 15, p. 100) is assumed to be negative.

Although MSC do not employ this approach, I next estimate a first-order vector autoregressive VAR(1) process for each multifactor model and its set of state variables, since this is how Campbell and Vuolteenaho (2004), Petkova (2006), and Koijen et al. (2010) estimate the innovations in the state variables:

$$z_{t+1} = a + \Gamma z_t + u_{t+1}, \tag{B.15}$$

where z_{t+1} is a (K + 1)x1 vector with the market excess return as its first element and the K state variables that are included in the respective multifactor model as its second to (K + 1)th element. In addition, a and Γ denote a (K + 1)x1 vector and a (K + 1)x(K + 1) matrix of constant parameters and u_{t+1} is a (K + 1)x1 vector of independent and identically distributed shocks. I define the vector of innovations, Δz_{t+1}, as the second to (K + 1)th element of the estimated shock vector, $\hat{u}_{t+1}$. The outcomes for the ICAPM specifications are shown in Table 53 (p. 182). Observe that the measures of the models' explanatory powers do not, again, differ that much from those obtained in the benchmark tests. Moreover, the results with regard to ICAPM criterion 1 are qualitatively similar to those of the original tests. Regarding the factor risk prices, one can see that γ_{TERM} is now negative within P (SM25) and positive within KLVN (SM25). Consequently, P no longer fulfills the ICAPM criteria 3a and 3b when tested over SM25.

Table 30. Multiple predictive regressions for state variables constructed from empirical factors (alternatives for SMB^* and HML^*).

This table shows the results for multiple predictive regressions with the monthly continuously compounded return on the European stock market as the left-hand side variable for horizons of one month (Panel A), 12 months (Panel B), and 60 months (Panel C). The right-hand side variables are the current European values of the size premium $CSMB$, the value premium $CHML$, the momentum premium $CUMD$, the liquidity factor CL, and the slopes of the yield curve $TERM$ and the corporate bond default spread DEF. The original sample period is November 1995 to December 2011 and one, 12, and 60 observations are lost in each of the one-, 12-, and 60-month horizon regressions, respectively. The first row of each regression shows the estimated slope coefficients and the second row displays the Newey–West t-ratios that were calculated with one, 12, or 60 lags (in parentheses). The levels of statistical significance of the estimated slopes are indicated by italics (10% level), an underline (5% level), and boldface (1% level). The R^2 (%) value presents the adjusted R^2 (%) value. Panel D shows the overall assessment of the variables' predictive characteristics, jointly considering the regression results for forecasting horizons of one month, three months, 12 months, 24 months, 36 months, 48 months, and 60 months. A + (-) means that the variable all in all forecasts positive (negative) market returns, while +/- indicates that the sign of the respective predictive slope is assessed as indeterminate. Parentheses indicate that the forecast is insignificant for all forecasting horizons.

Row	CSMB	CHML	CUMD	CL	TERM	DEF	R^2 (%)
Panel A: q = 1							
1	0.01	-0.00					-0.71
	(0.78)	(-0.43)					
2	0.02	-0.00	0.01				-1.14
	(0.94)	(-0.51)	(0.49)				
3	0.03	-0.02		0.78			1.53
	(1.54)	(*-1.87*)		(<u>2.16</u>)			
4	0.03	-0.00			1.46	-0.63	7.38
	(*1.69*)	(-0.55)			(**3.70**)	(-0.62)	
Panel B: q = 12							
1	-0.03	0.02					-1.04
	(-0.10)	(0.16)					
2	0.04	0.01	0.22				0.98
	(0.14)	(0.10)	(1.24)				
3	0.05	-0.07		4.76			2.35
	(0.14)	(-0.55)		(1.20)			
4	0.09	0.03			13.78	6.00	31.94
	(0.37)	(0.32)			(**4.27**)	(0.91)	
Panel C: q = 60							
1	-0.63	0.28					22.41
	(<u>-2.07</u>)	(*1.86*)					
2	-0.62	0.31	0.20				22.57
	(<u>-2.22</u>)	(1.55)	(0.40)				
3	-0.61	0.30		-3.46			22.05
	(<u>-1.96</u>)	(<u>2.06</u>)		(-1.02)			
4	-0.42	0.14			15.24	50.41	46.64
	(<u>-2.48</u>)	(1.21)			(<u>2.04</u>)	(**4.75**)	
Panel D: Overall assessment							
1	-	+					
2	-	(+/-)	(+/-)				
3	-	+/-		+			
4	+/-	(+/-)			+	+	

Table 31. Multiple predictive regressions for state variables constructed from empirical factors ($SVAR$, alternatives for SMB* and HML*).

This table shows the results for multiple predictive regressions with the variance ($SVAR$) of the European stock market as the left-hand side variable for horizons of one month (Panel A), 12 months (Panel B), and 60 months (Panel C). The right-hand side variables are the current European values of the size premium $CSMB$, the value premium $CHML$, the momentum premium $CUMD$, the liquidity factor CL, the slope of the yield curve $TERM$, and the corporate bond default spread DEF. The original sample period is November 1995 to December 2011 and one, 12, and 60 observations are lost in each of the one-, 12-, and 60-month horizon regressions, respectively. The first row of each regression shows the estimated slope coefficients and the second row displays the Newey–West t-ratios that were calculated with one, 12, or 60 lags (in parentheses). The levels of statistical significance of the estimated slopes are indicated by italics (10% level), an underline (5% level), and boldface (1% level). The R^2 (%) value presents the adjusted R^2 (%) value. Panel D shows the overall assessment of the variables' predictive characteristics, jointly considering the regression results for forecasting horizons of one month, three months, 12 months, 24 months, 36 months, 48 months, and 60 months. A + (-) means that the variable all in all forecasts positive (negative) market returns, while +/- indicates that the sign of the respective predictive slope is assessed as indeterminate. Parentheses indicate that the forecast is insignificant for all forecasting horizons.

Row	CSMB	CHML	CUMD	CL	TERM	DEF	R^2 (%)
Panel A: q = 1							
1	-0.00	-0.00					1.61
	(-1.10)	(*-1.88*)					
2	0.00	-0.00	0.00				4.26
	(0.08)	(-2.21)	(*1.94*)				
3	-0.00	0.00		-0.09			5.75
	(-2.50)	(0.84)		(-2.34)			
4	-0.00	-0.00			-0.12	0.28	21.40
	(-2.53)	(-1.46)			(-2.00)	(**3.42**)	
Panel B: q = 12							
1	0.00	-0.02					6.16
	(0.08)	(-1.42)					
2	0.00	-0.02	-0.00				5.61
	(0.05)	(-1.41)	(-0.10)				
3	-0.01	-0.00		-0.86			14.89
	(-0.44)	(-0.07)		(-1.52)			
4	-0.02	-0.01			-1.43	1.06	39.35
	(-0.66)	(*-1.77*)			(**-2.77**)	(1.53)	
Panel C: q = 60							
1	0.18	-0.04					50.93
	(**7.19**)	(**-3.09**)					
2	0.18	-0.05	-0.07				53.86
	(**7.16**)	(**-3.08**)	(*-1.91*)				
3	0.19	-0.04		-0.54			50.77
	(**6.49**)	(**-2.80**)		(-0.63)			
4	0.14	-0.02			-1.82	-9.29	70.11
	(**16.51**)	(-1.22)			(**-2.96**)	(**-3.32**)	
Panel D: Overall assessment							
1	+	-					
2	+	-	+/-				
3	+/-	-		-			
4	+/-	-			-	+/-	

Table 32. Multiple predictive regressions for ICAPM state variables (market variance; Beeler and Campbell, 2012).

This table shows the results for multiple predictive regressions with the variance of the European stock market, which is obtained using the proxy suggested by Beeler and Campbell (2012), as the left-hand side variable for horizons of one month (Panel A), 12 months (Panel B), and 60 months (Panel C). The right-hand side variables are the current European values of the slope of the yield curve $TERM$, the corporate bond default spread DEF, the market dividend-to-price ratio DY, the short-term risk-free rate RF, the aggregate price–earnings ratio PE, the value spread VS, and the Cochrane–Piazzesi factor CP. The original sample period is December 1990 to December 2011. However, one, 12, and 60 observations are lost in each of the one-, 12-, and 60-month horizon regressions, respectively. The first row of each regression shows the estimated slope coefficients and the second row displays the Newey–West t-ratios that were calculated with one, 12, or 60 lags (in parentheses). The levels of statistical significance of the estimated slopes are indicated by italics (10% level), an underline (5% level), and boldface (1% level). The R^2 (%) value presents the adjusted R^2 (%). Panel D shows the overall assessment of the variables' predictive characteristics, jointly considering the regression results for forecasting horizons of one month, three months, 12 months, 24 months, 36 months, 48 months, and 60 months. A + (-) means that the variable all in all forecasts positive (negative) market returns, while +/- indicates that the sign of the respective predictive slope is assessed as indeterminate. Parentheses indicate that the forecast is insignificant for all forecasting horizons.

Row	TERM	DEF	DY	RF	PE	VS	CP	R^2 (%)
Panel A: q = 1								
1	-0.01	0.17						4.84
	(-0.98)	(*1.94*)						
2	-0.01	0.19	-0.00	0.00				3.80
	(-0.46)	(*1.70*)	(-0.69)	(0.05)				
3	-0.02				-0.00	0.00		0.65
	(-1.45)				(-0.79)	(*1.78*)		
4	-0.03						0.03	0.05
	(-1.54)						(0.96)	
Panel B: q=12								
1	-0.02	0.10						13.06
	(-1.07)	(**2.58**)						
2	-0.00	0.25	-0.00	0.02				29.65
	(-0.08)	(**3.65**)	(**-3.17**)	(1.19)				
3	-0.02				0.00	0.00		5.41
	(-1.19)				(1.00)	(0.06)		
4	-0.02						-0.01	2.66
	(-0.75)						(-0.29)	
Panel C: q=60								
1	0.01	-0.10						31.18
	(*1.86*)	(**-2.60**)						
2	-0.00	-0.10	-0.00	-0.01				71.15
	(-0.41)	(**-4.32**)	(**-5.68**)	(-1.47)				
3	0.01				0.00	-0.00		66.89
	(<u>2.30</u>)				(**9.48**)	(**-2.60**)		
4	0.02						-0.00	20.18
	(<u>1.99</u>)						(-0.18)	
Panel D: Overall assessment								
1	+	+/-						
2	(-)	+/-	-	(+/-)				
3	+				+	+/-		
4	+						(+/-)	

Table 33. Multiple predictive regressions for state variables constructed from empirical factors (market variance; Beeler and Campbell, 2012).

This table shows the results for multiple predictive regressions with the variance of the European stock market, which is obtained using the proxy suggested by Beeler and Campbell (2012), as the left-hand side variable, for horizons of one month (Panel A), 12 months (Panel B), and 60 months (Panel C). The right-hand side variables are the current European values of the size premium SMB^*, the value premium HML^*, the momentum premium $CUMD$, the liquidity factor CL, the slope of the yield curve $TERM$, and the corporate bond default spread DEF. The original sample period is November 1990 to December 2011. However, one, 12, and 60 observations are lost in each of the one-, 12-, and 60-month horizon regressions, respectively, and the original sample period for the regressions containing $CUMD$ or CL is from November 1995 to December 2011. The first row of each regression shows the estimated slope coefficients and the second row displays the Newey–West t-ratios that were calculated with one, 12, or 60 lags (in parentheses). The levels of statistical significance of the estimated slopes are indicated by italics (10% level), an underline (5% level), and boldface (1% level). The R^2 (%) value presents the adjusted R^2 (%). Panel D shows the overall assessment of the variables' predictive characteristics, jointly considering the regression results for forecasting horizons of one month, three months, 12 months, 24 months, 36 months, 48 months, and 60 months. A + (-) means that the variable all in all forecasts positive (negative) market returns, while +/- indicates that the sign of the respective predictive slope is assessed as indeterminate. Parentheses indicate that the forecast is insignificant for all forecasting horizons.

Row	SMB*	HML*	CUMD	CL	TERM	DEF	R² (%)
Panel A: q = 1							
1	-0.00	0.00					2.63
	(-2.56)	(0.20)					
2	-0.00	0.00	0.00				4.93
	(-2.19)	(*1.66*)	(2.36)				
3	-0.00	0.00		-0.03			3.38
	(**-2.72**)	(0.30)		(-2.13)			
4	-0.00	-0.00			-0.01	0.18	7.80
	(*-1.71*)	(-1.05)			(-0.83)	(2.03)	
Panel B: q = 12							
1	-0.00	-0.00					15.75
	(-2.29)	(-0.84)					
2	-0.00	0.00	-0.00				11.25
	(-2.19)	(0.24)	(-0.12)				
3	-0.00	-0.00		-0.05			25.50
	(-2.32)	(-0.57)		(*-1.87*)			
4	-0.00	-0.00			-0.02	0.11	33.61
	(*-1.86*)	(*-1.92*)			(-1.16)	(**4.04**)	
Panel C: q = 60							
1	0.00	-0.00					21.68
	(0.60)	(**-2.77**)					
2	0.00	-0.00	-0.00				18.52
	(0.89)	(**-3.13**)	(**-3.96**)				
3	0.00	-0.00		-0.01			11.61
	(0.64)	(-2.30)		(-1.15)			
4	-0.00	-0.00			0.01	-0.12	57.15
	(-1.10)	(-2.54)			(**2.93**)	(**-3.15**)	
Panel D: Overall assessment							
1	-	-					
2	-	+/-	+/-				
3	+/-	-		-			
4	-	-			+	+/-	

Table 34. Multiple predictive regressions for ICAPM state variables (market variance; Bansal et al., 2005).

This table shows the results for multiple predictive regressions with the variance of the European stock market, which is obtained using the proxy suggested by Bansal et al. (2005), as the left-hand side variable, for horizons of one month (Panel A), 12 months (Panel B), and 60 months (Panel C). The right-hand side variables are the current European values of the slope of the yield curve $TERM$, the corporate bond default spread DEF, the market dividend-to-price ratio DY, the short-term risk-free rate RF, the aggregate price–earnings ratio PE, the value spread VS, and the Cochrane–Piazzesi factor CP. The original sample period is December 1990 to December 2011. However, one, 12, and 60 observations are lost in each of the one-, 12-, and 60-month horizon regressions, respectively. The first row of each regression shows the estimated slope coefficients and the second row displays the Newey–West t-ratios that were calculated with one, 12, or 60 lags (in parentheses). The levels of statistical significance of the estimated slopes are indicated by italics (10% level), an underline (5% level), and boldface (1% level). The R^2 (%) value presents the adjusted R^2 (%). Panel D shows the overall assessment of the variables' predictive characteristics, jointly considering the regression results for forecasting horizons of one month, three months, 12 months, 24 months, 36 months, 48 months, and 60 months. A + (-) means that the variable all in all forecasts positive (negative) market returns, while +/- indicates that the sign of the respective predictive slope is assessed as indeterminate. Parentheses indicate that the forecast is insignificant for all forecasting horizons.

Row	TERM	DEF	DY	RF	PE	VS	CP	R^2 (%)
Panel A: q = 1								
1	-1.29	33.34						1.34
	(-0.28)	(**2.99**)						
2	-6.40	38.74	-0.09	-3.16				1.14
	(-0.79)	(*1.93*)	(-0.69)	(-0.64)				
3	-1.97				-0.14	0.90		0.70
	(-0.41)				(-0.80)	($\underline{2.14}$)		
4	-2.13						-1.00	-0.68
	(-0.31)						(-0.10)	
Panel B: q=12								
1	-4.83	21.06						9.89
	(-1.09)	($\underline{2.42}$)						
2	-2.73	63.11	-0.35	4.13				26.92
	(-0.33)	(**4.42**)	(**-4.43**)	(0.85)				
3	-5.38				0.13	0.47		10.66
	(-1.12)				(0.88)	(1.33)		
4	-0.58						-11.83	7.39
	(-0.11)						(*-1.95*)	
Panel C: q=60								
1	3.37	-38.82						41.01
	($\underline{1.97}$)	(**-3.99**)						
2	0.84	-30.63	-0.16	0.01				69.34
	(0.64)	(**-5.26**)	(**-5.81**)	(0.01)				
3	2.39				0.46	-0.33		59.97
	(*1.84*)				(**7.00**)	($\underline{-2.29}$)		
4	5.68						-2.56	17.62
	($\underline{2.03}$)						(-0.90)	
Panel D: Overall assessment								
1	+	+/-						
2	(+/-)	+/-	-	+				
3	+				+	+/-		
4	+						-	

Table 35. Multiple predictive regressions for state variables constructed from empirical factors (market variance; Bansal et al., 2005).

This table shows the results for multiple predictive regressions with the variance of the European stock market, which is obtained using the proxy suggested by Bansal et al. (2005), as the left-hand side variable, for horizons of one month (Panel A), 12 months (Panel B), and 60 months (Panel C). The right-hand side variables are the current European values of the size premium SMB^*, the value premium HML^*, the momentum premium $CUMD$, the liquidity factor CL, the slope of the yield curve $TERM$, and the corporate bond default spread DEF. The original sample period is November 1990 to December 2011. However, one, 12, and 60 observations are lost in each of the one-, 12-, and 60-month horizon regressions, respectively, and the original sample period for the regressions containing $CUMD$ or CL is from November 1995 to December 2011. The first row of each regression shows the estimated slope coefficients and the second row displays the Newey–West t-ratios that were calculated with one, 12, or 60 lags (in parentheses). The levels of statistical significance of the estimated slopes are indicated by italics (10% level), an underline (5% level), and boldface (1% level). The R^2 (%) value presents the adjusted R^2 (%). Panel D shows the overall assessment of the variables' predictive characteristics, jointly considering the regression results for forecasting horizons of one month, three months, 12 months, 24 months, 36 months, 48 months, and 60 months. A + (-) means that the variable all in all forecasts positive (negative) market returns, while +/- indicates that the sign of the respective predictive slope is assessed as indeterminate. Parentheses indicate that the forecast is insignificant for all forecasting horizons.

Row	SMB*	HML*	CUMD	CL	TERM	DEF	R² (%)
Panel A: q = 1							
1	-0.17	-0.05					0.30
	(-0.83)	(-0.84)					
2	-0.26	0.02	0.35				-0.05
	(-1.09)	(0.31)	(0.74)				
3	-0.31	-0.03		-7.86			0.75
	(-1.31)	(-0.41)		(*-1.68*)			
4	-0.06	-0.12			-1.41	41.58	2.38
	(-0.26)	(*-1.79*)			(-0.31)	(**3.48**)	
Panel B: q = 12							
1	-0.31	-0.07					20.27
	(<u>-2.39</u>)	(<u>-2.03</u>)					
2	-0.36	-0.04	-0.25				19.03
	(<u>-2.51</u>)	(-0.88)	(-0.81)				
3	-0.39	-0.06		-13.65			36.66
	(**-2.58**)	(<u>-2.32</u>)		(**-2.70**)			
4	-0.22	-0.12			-4.83	28.48	34.19
	(*-1.79*)	(**-3.02**)			(-1.25)	(**4.16**)	
Panel C: q = 60							
1	0.04	-0.07					28.11
	(0.52)	(**-3.59**)					
2	0.03	-0.08	-0.33				29.20
	(0.49)	(**-4.17**)	(**-3.65**)				
3	0.04	-0.06		-6.26			25.27
	(0.48)	(**-3.29**)		(<u>-2.02</u>)			
4	-0.08	-0.04			2.96	-42.53	63.07
	(-1.44)	(**-2.99**)			(**2.65**)	(**-4.46**)	
Panel D: Overall assessment							
1	-	-					
2	-	-	-				
3	-	-		-			
4	-	-			+	+/-	

Table 36. Multiple predictive regressions for ICAPM state variables (squared return).
This table shows the results for multiple predictive regressions with the monthly squared continuously compounded return on the European stock market as the left-hand side variable for horizons of one month (Panel A), 12 months (Panel B), and 60 months (Panel C). The right-hand side variables are the current European values of the slope of the yield curve $TERM$, the corporate bond default spread DEF, the market dividend-to-price ratio DY, the short-term risk-free rate RF, the aggregate price–earnings ratio PE, the value spread VS, and the Cochrane–Piazzesi factor CP. The original sample period is December 1990 to December 2011. However, one, 12, and 60 observations are lost in each of the one-, 12-, and 60-month horizon regressions, respectively. The first row of each regression shows the estimated slope coefficients and the second row displays the Newey–West t-ratios that were calculated with one, 12, or 60 lags (in parentheses). The levels of statistical significance of the estimated slopes are indicated by italics (10% level), an underline (5% level), and boldface (1% level). The R^2 (%) value presents the adjusted R^2 (%). Panel D shows the overall assessment of the variables' predictive characteristics, jointly considering the regression results for forecasting horizons of one month, three months, 12 months, 24 months, 36 months, 48 months, and 60 months. A + (-) means that the variable all in all forecasts positive (negative) market returns, while +/- indicates that the sign of the respective predictive slope is assessed as indeterminate. Parentheses indicate that the forecast is insignificant for all forecasting horizons.

Row	TERM	DEF	DY	RF	PE	VS	CP	R² (%)
Panel A: q = 1								
1	-0.00	0.17						5.07
	(-0.23)	(2.13)						
2	-0.00	0.20	-0.00	-0.00				4.49
	(-0.18)	(*1.92*)	(-0.62)	(0.06)				
3	-0.01				-0.00	0.00		0.47
	(-0.57)				(-0.82)	(*1.87*)		
4	-0.01						0.01	-0.62
	(-0.63)						(0.27)	
Panel B: q=12								
1	0.67	2.02						4.10
	(0.93)	(2.06)						
2	0.71	1.59	0.00	0.00				3.43
	(0.56)	(0.70)	(0.25)	(0.00)				
3	0.57				0.00	-0.05		2.76
	(0.81)				(0.09)	(-1.08)		
4	1.44						-1.99	6.39
	(*1.79*)						(*-1.78*)	
Panel C: q=60								
1	0.35	-68.87						4.69
	(0.05)	(-0.92)						
2	30.41	-61.54	0.49	16.28				57.95
	(**3.63**)	(-0.90)	(**3.82**)	(2.07)				
3	9.61				-1.32	1.20		22.28
	(0.99)				(-2.32)	(1.33)		
4	8.98						-15.48	2.19
	(1.09)						(-2.02)	
Panel D: Overall assessment								
1	+	+						
2	+	+/-	+	+				
3	+				-	+		
4	+						-	

Table 37. Multiple predictive regressions for state variables constructed from empirical factors (squared return).

This table shows the results for multiple predictive regressions with the monthly squared continuously compounded return on the European stock market as the left-hand side variable for horizons of one month (Panel A), 12 months (Panel B), and 60 months (Panel C). The right-hand side variables are the current European values of the size premium SMB^*, the value premium HML^*, the momentum premium $CUMD$, the liquidity factor CL, the slope of the yield curve $TERM$, and the corporate bond default spread DEF. The original sample period is November 1990 to December 2011. However, one, 12, and 60 observations are lost in each of the one-, 12-, and 60-month horizon regressions, respectively, and the original sample period for the regressions containing $CUMD$ or CL is from November 1995 to December 2011. The first row of each regression shows the estimated slope coefficients and the second row displays the Newey–West t-ratios that were calculated with one, 12, or 60 lags (in parentheses). The levels of statistical significance of the estimated slopes are indicated by italics (10% level), an underline (5% level), and boldface (1% level). The R^2 (%) value presents the adjusted R^2 (%). Panel D shows the overall assessment of the variables' predictive characteristics, jointly considering the regression results for forecasting horizons of one month, three months, 12 months, 24 months, 36 months, 48 months, and 60 months. A + (-) means that the variable all in all forecasts positive (negative) market returns, while +/- indicates that the sign of the respective predictive slope is assessed as indeterminate. Parentheses indicate that the forecast is insignificant for all forecasting horizons.

Row	SMB*	HML*	CUMD	CL	TERM	DEF	R^2 (%)
Panel A: q = 1							
1	-0.00	-0.00					1.85
	(*-1.91*)	(-0.14)					
2	-0.00	0.00	0.00				4.78
	(-1.46)	(1.32)	(**2.59**)				
3	-0.00	0.00		-0.03			1.34
	(*-1.85*)	(0.01)		(*-1.68*)			
4	-0.00	-0.00			-0.00	0.19	7.45
	(-1.09)	(-1.35)			(-0.16)	(2.28)	
Panel B: q = 12							
1	-0.03	0.00					2.83
	(*-1.86*)	(1.07)					
2	-0.03	0.01	-0.01				4.61
	(-2.50)	(*1.67*)	(-0.09)				
3	-0.03	0.01		-0.05			4.59
	(-2.44)	(2.28)		(-0.07)			
4	-0.03	0.00			0.72	1.64	6.46
	(-2.06)	(0.89)			(0.99)	(1.49)	
Panel C: q = 60							
1	0.17	0.24					30.00
	(0.85)	(**3.12**)					
2	-0.17	0.12	0.66				30.23
	(-1.07)	(**4.73**)	(2.48)				
3	-0.13	0.11		-8.55			22.08
	(-0.71)	(**4.24**)		(-1.09)			
4	-0.07	0.30			2.43	-93.38	38.17
	(-0.31)	(**2.74**)			(0.52)	(-1.44)	
Panel D: Overall assessment							
1	-	+					
2	-	+	+				
3	-	+		+/-			
4	-	+			+	+/-	

Table 38. Factor risk premiums for ICAPM specifications (intercept).
This table reports the estimation results for the multifactor models with economically motivated factors using first-stage GMM. Panel A displays the results for the 25 size/book-to-market portfolios of the European stock market as test assets and Panel B presents the results for the 25 size/momentum portfolios. I add the market excess return to both sets of test assets. The market (covariance) risk price is represented by γ and γ_{TERM}, γ_{DEF}, γ_{DY}, γ_{RF}, γ_{PE}, γ_{VS}, and γ_{CP} denote the (covariance) risk prices associated with the slope of the yield curve, the corporate bond default spread, the market dividend-to-price ratio, the short-term risk-free rate, the aggregate price–earnings ratio, the value spread, and the Cochrane–Piazzesi factor, respectively. The term γ_0 represents the intercept. The estimates of the (covariance) risk prices are reported in the first row that corresponds to each model and the second row contains the asymptotic GMM robust t-statistics (in parentheses). The levels of statistical significance of the estimated (covariance) risk prices are indicated by italics (10% level), an underline (5% level), and boldface (1% level). The second to last column shows the average absolute pricing errors (%) and the last column presents the OLS cross-sectional R^2 value of each model. The sample period is December 1990 to December 2011.

Model	γ_0	γ	γ_{TERM}	γ_{DEF}	γ_{DY}	γ_{RF}	γ_{PE}	γ_{VS}	γ_{CP}	MAE (%)	R^2_{OLS}
Panel A: SBM25											
CAPM	0.01	-1.81								0.22	0.02
	(<u>2.30</u>)	(-0.98)									
HL	0.01	-2.64	-147.33	49.95						0.21	0.15
	(<u>2.49</u>)	(-0.92)	(-0.77)	(0.24)							
P	0.01	-27.92	220.06	784.26	-9.14	-68.18				0.11	0.68
	(**4.38**)	(**-4.69**)	(<u>2.11</u>)	(**4.38**)	(**-5.89**)	(-0.45)					
CV	0.01	85.46	-218.69				-81.46	-15.92		0.10	0.81
	(**4.04**)	(**2.84**)	(**-4.26**)				(**-2.85**)	(**-7.52**)			
KLVN	0.01	-3.29	-189.86						10.22	0.21	0.14
	(*1.72*)	(-0.76)	(-1.00)						(0.08)		
Panel B: SM25											
CAPM	0.02	-6.69								0.24	0.32
	(**5.15**)	(**-3.90**)									
HL	0.02	-1.92	210.04	572.63						0.21	0.57
	(**6.59**)	(-0.98)	(1.59)	(**2.81**)							
P	0.01	20.13	92.44	67.91	6.33	258.42				0.16	0.67
	(<u>2.24</u>)	(**3.86**)	(0.42)	(0.23)	(**4.90**)	(1.01)					
CV	0.02	77.20	-133.98				-77.74	4.10		0.19	0.60
	(**6.75**)	(*1.93*)	(-1.40)				(<u>-2.08</u>)	(0.89)			
KLVN	0.02	-12.39	-330.27						395.92	0.23	0.49
	(**6.16**)	(**-4.40**)	(<u>-2.37</u>)						(<u>2.43</u>)		

Table 39. Factor risk premiums for empirical risk factors (intercept).

This table reports the estimation results for the multifactor models with empirical factors using first-stage GMM. Panel A displays the results for the 25 size/book-to-market portfolios of the European stock market as test assets and Panel B presents the results for the 25 size/momentum portfolios. I add the market excess return to both sets of test assets. The market (covariance) risk price is represented by γ and γ_{SMB}, γ_{HML}, γ_{UMD}, γ_L, γ_{TERM}, and γ_{DEF} denote the (covariance) risk prices associated with the size factor, the value factor, the momentum factor, the liquidity factor, the slope of the yield curve, and the corporate bond default spread. The term γ_0 represents the intercept. The estimates of the (covariance) risk prices are reported in the first row that corresponds to each model and the second row contains the asymptotic GMM robust t-statistics (in parentheses). The levels of statistical significance of the estimated (covariance) risk prices are indicated by italics (10% level), an underline (5% level), and boldface (1% level). The second to last column shows the average absolute pricing errors (%) and the last column presents the OLS cross-sectional R^2 value of each model. The sample period is December 1990 to December 2011.

Model	γ_0	γ	γ_{SMB}	γ_{HML}	γ_{UMD}	γ_L	γ_{TERM}	γ_{DEF}	MAE (%)	R^2_{OLS}
Panel A: SBM25										
FF3	0.01	-4.92	-5.70	6.83					0.12	0.69
	(**3.56**)	(-2.57)	(**-3.72**)	(**7.20**)						
C	0.01	2.44	-3.15	9.99	10.40				0.09	0.77
	(2.48)	(0.86)	(**-3.14**)	(**6.92**)	(2.33)					
PS	0.01	-0.82	-2.95	9.67		-276.58			0.11	0.71
	(**4.06**)	(-0.47)	(*-1.72*)	(**4.65**)		(-1.63)				
FF5	0.01	-1.34	-3.01	8.16			140.83	321.91	0.11	0.74
	(**4.64**)	(-0.76)	(-0.76)	(**6.52**)			(0.98)	(*1.97*)		
Panel B: SM25										
FF3	0.02	-7.15	-4.45	-13.30					0.20	0.55
	(**6.12**)	(**-3.96**)	(-1.26)	(**-2.62**)						
C	0.01	1.55	-1.03	3.64	5.51				0.16	0.65
	(0.98)	(0.56)	(-0.36)	(0.44)	(**3.20**)					
PS	0.02	-7.28	-4.61	-13.33		10.25			0.20	0.55
	(**6.04**)	(**-2.17**)	(-0.95)	(**-2.72**)		(0.05)				
FF5	0.02	1.89	11.64	-7.44			-151.81	561.92	0.20	0.60
	(**3.93**)	(0.36)	(0.93)	(-1.30)			(-0.58)	(*1.82*)		

Table 40. Factor risk premiums for ICAPM specifications (second-stage GMM).

This table reports the estimation results for the multifactor models with economically motivated factors using second-stage GMM. Panel A displays the results for the 25 size/book-to-market portfolios of the European stock market as test assets and Panel B presents the results for the 25 size/momentum portfolios. I add the market excess return to both sets of test assets. The market (covariance) risk price is represented by γ and γ_{TERM}, γ_{DEF}, γ_{DY}, γ_{RF}, γ_{PE}, γ_{VS}, and γ_{CP} denote the (covariance) risk prices associated with the slope of the yield curve, the corporate bond default spread, the market dividend-to-price ratio, the short-term risk-free rate, the aggregate price–earnings ratio, the value spread, and the Cochrane–Piazzesi factor. The estimates of the (covariance) risk prices are reported in the first row that corresponds to each model and the second row contains the asymptotic GMM robust t-statistics (in parentheses). The levels of statistical significance of the estimated (covariance) risk prices are indicated by italics (10% level), an underline (5% level), and boldface (1% level). The last column presents the WLS cross-sectional R^2 value of each model. The sample period is December 1990 to December 2011.

Model	γ	γ_{TERM}	γ_{DEF}	γ_{DY}	γ_{RF}	γ_{PE}	γ_{VS}	γ_{CP}	R^2_{WLS}
Panel A: SBM25									
CAPM	5.94								-0.07
	(**3.98**)								
HL	8.54	168.03	355.56						-0.04
	(**4.83**)	(2.01)	(**3.37**)						
P	4.70	68.15	152.46	-0.21	-117.29				0.44
	(0.49)	(0.45)	(0.61)	(-0.07)	(-0.73)				
CV	70.48	32.11				-61.12	-18.82		0.67
	(**2.91**)	(0.27)				(**-2.72**)	(**-3.46**)		
KLVN	7.23	102.04						-109.83	0.00
	(**4.32**)	(1.16)						(-2.12)	
Panel B: SM25									
CAPM	5.68								-0.22
	(**3.82**)								
HL	10.65	360.99	662.32						-0.03
	(**4.59**)	(**3.91**)	(**3.25**)						
P	47.79	338.59	-42.93	14.00	460.77				0.60
	(**5.51**)	(**3.47**)	(-0.34)	(**5.03**)	(**3.39**)				
CV	90.44	142.98				-81.47	-2.43		0.13
	(**5.16**)	(*1.76*)				(**-5.04**)	(-0.68)		
KLVN	3.55	88.90						51.06	-0.17
	(1.98)	(1.22)						(0.67)	

Table 41. Factor risk premiums for empirical risk factors (second-stage GMM).

This table reports the estimation results for the multifactor models with empirical factors using second-stage GMM. Panel A displays the results for the 25 size/book-to-market portfolios of the European stock market as test assets and Panel B presents the results for the 25 size/momentum portfolios. I add the market excess return to both sets of test assets. The market (covariance) risk price is represented by γ and γ_{SMB}, γ_{HML}, γ_{UMD}, γ_{L}, γ_{TERM}, and γ_{DEF} denote the (covariance) risk prices associated with the size factor, the value factor, the momentum factor, the liquidity factor, the slope of the yield curve, and the corporate bond default spread. The estimates of the (covariance) risk prices are reported in the first row that corresponds to each model and the second row contains the asymptotic GMM robust t-statistics (in parentheses). The levels of statistical significance of the estimated (covariance) risk prices are indicated by italics (10% level), an underline (5% level), and boldface (1% level). The second to last column shows the average absolute pricing errors (%) and the last column presents the WLS cross-sectional R^2 value of each model. The sample period is December 1990 to December 2011.

Model	γ	γ_{SMB}	γ_{HML}	γ_{UMD}	γ_{L}	γ_{TERM}	γ_{DEF}	R^2_{WLS}
Panel A: SBM25								
FF3	3.79	-4.54	13.08					0.50
	(**2.52**)	(-1.60)	(**5.69**)					
C	10.55	-4.07	12.84	25.87				0.72
	(**4.97**)	(-1.34)	(**4.50**)	(**6.67**)				
PS	12.34	5.77	3.97		-420.83			0.56
	(**5.16**)	(1.43)	(1.12)		(**-3.25**)			
FF5	10.91	3.81	14.53			181.64	636.52	0.57
	(**4.36**)	(0.92)	(**5.34**)			(*1.84*)	(**3.76**)	
Panel B: SM25								
FF3	7.91	5.49	-25.64					0.02
	(**4.75**)	(*1.80*)	(**-6.37**)					
C	8.66	-1.03	8.42	10.13				0.56
	(**4.85**)	(-0.35)	(<u>2.30</u>)	(**5.09**)				
PS	9.77	4.20	-14.85		-400.84			0.03
	(**4.30**)	(1.19)	(**-3.47**)		(**-3.82**)			
FF5	16.04	24.39	2.45			-148.25	865.94	0.41
	(<u>2.42</u>)	(*1.89*)	(0.26)			(-0.76)	(<u>1.99</u>)	

Table 42. Factor risk premiums for ICAPM specifications (excess bond returns).

This table reports the estimation results for the multifactor models with economically motivated factors using first-stage GMM. Panel A displays the results for the 25 size/book-to-market portfolios of the European stock market as test assets and Panel B presents the results for the 25 size/momentum portfolios. I add the market excess return as well as the excess returns on seven German government bonds to both sets of test assets. The market (covariance) risk price is represented by γ and γ_{TERM}, γ_{DEF}, γ_{DY}, γ_{RF}, γ_{PE}, γ_{VS}, and γ_{CP} denote the (covariance) risk prices associated with the slope of the yield curve, the corporate bond default spread, the market dividend-to-price ratio, the short-term risk-free rate, the aggregate price–earnings ratio, the value spread, and the Cochrane–Piazzesi factor. The estimates of the (covariance) risk prices are reported in the first row that corresponds to each model and the second row contains the asymptotic GMM robust t-statistics (in parentheses). The levels of statistical significance of the estimated (covariance) risk prices are indicated by italics (10% level), an underline (5% level), and boldface (1% level). The second to last column shows the average absolute pricing errors (%) and the last column presents the OLS cross-sectional R^2 value of each model. The sample period is December 1990 to December 2011.

Model	γ	γ_{TERM}	γ_{DEF}	γ_{DY}	γ_{RF}	γ_{PE}	γ_{VS}	γ_{CP}	MAE (%)	R^2_{OLS}
Panel A: SBM25										
CAPM	2.31								0.26	-0.23
	(1.49)									
HL	2.87	-114.83	-0.96						0.20	0.14
	(1.41)	(<u>-2.48</u>)	(-0.01)							
P	-12.80	14.88	454.20	-6.17	-410.69				0.13	0.57
	(-1.18)	(0.24)	(*1.78*)	(*-1.71*)	(**-2.92**)					
CV	75.76	-103.53				-68.01	-15.94		0.10	0.74
	(<u>2.31</u>)	(*-1.83*)				(<u>-2.21</u>)	(*-1.76*)			
KLVN	3.83	-9.04						-116.25	0.17	0.19
	(<u>2.06</u>)	(-0.12)						(*-1.82*)		
Panel B: SM25										
CAPM	2.36								0.35	-0.29
	(1.51)									
HL	4.91	-62.60	200.05						0.30	-0.12
	(<u>2.31</u>)	(-1.27)	(1.41)							
P	30.55	-162.46	-339.95	9.22	34.92				0.17	0.65
	(**2.70**)	(**-2.85**)	(-1.44)	(<u>2.42</u>)	(0.27)					
CV	38.20	-121.29				-33.22	12.78		0.26	0.22
	(1.43)	(<u>-2.34</u>)				(-1.36)	(<u>2.01</u>)			
KLVN	3.25	-53.56						-48.94	0.29	-0.17
	(*1.74*)	(-0.70)						(-0.79)		

Table 43. Factor risk premiums for empirical risk factors (excess bond returns).

This table reports the estimation results for the multifactor models with empirical factors using first-stage GMM. Panel A displays the results for the 25 size/book-to-market portfolios of the European stock market as test assets and Panel B presents the results for the 25 size/momentum portfolios. I add the market excess return as well as the excess returns on seven German government bonds to both sets of test assets. The market (covariance) risk price is represented by γ and γ_{SMB}, γ_{HML}, γ_{UMD}, γ_L, γ_{TERM}, and γ_{DEF} denote the (covariance) risk prices associated with the size factor, the value factor, the momentum factor, the liquidity factor, the slope of the yield curve, and the corporate bond default spread. The estimates of the (covariance) risk prices are reported in the first row that corresponds to each model and the second row contains the asymptotic GMM robust t-statistics (in parentheses). The levels of statistical significance of the estimated (covariance) risk prices are indicated by italics (10% level), an underline (5% level), and boldface (1% level). The second to last column shows the average absolute pricing errors (%) and the last column presents the OLS cross-sectional R^2 value of each model. The sample period is December 1990 to December 2011.

Model	γ	γ_{SMB}	γ_{HML}	γ_{UMD}	γ_L	γ_{TERM}	γ_{DEF}	MAE (%)	R^2_{OLS}
Panel A: SBM25									
FF3	1.86	-4.43	7.08					0.17	0.29
	(1.19)	(-1.45)	(**2.99**)						
C	7.12	-2.86	10.96	13.84				0.15	0.48
	(**2.96**)	(-0.84)	(**3.57**)	(**2.92**)					
PS	11.11	3.59	16.10		-895.18			0.14	0.57
	(<u>2.37</u>)	(0.51)	(*1.82*)		(<u>-2.12</u>)				
FF5	7.43	6.95	9.85			-79.99	445.46	0.12	0.64
	(<u>2.43</u>)	(1.56)	(**2.86**)			(-1.54)	(<u>2.20</u>)		
Panel B: SM25									
FF3	3.26	0.37	-13.94					0.33	-0.12
	(*1.80*)	(0.11)	(*-1.89*)						
C	4.44	-1.58	7.93	7.03				0.19	0.49
	(<u>2.51</u>)	(-0.49)	(*1.88*)	(**3.45**)					
PS	10.98	11.59	-11.53		-726.23			0.32	-0.02
	(**2.74**)	(<u>2.23</u>)	(-1.03)		(**-3.63**)				
FF5	15.21	23.91	0.55			-59.37	1069.83	0.23	0.29
	(**2.67**)	(**2.59**)	(0.07)			(-0.74)	(*1.86*)		

Table 44. Factor risk premiums for ICAPM specifications (excluding excess market return).

This table reports the estimation results for the multifactor models with economically motivated factors using first-stage GMM. Panel A displays the results for the 25 size/book-to-market portfolios of the European stock market as test assets and Panel B presents the results for the 25 size/momentum portfolios. I do not add the market excess return to the test assets. The market (covariance) risk price is represented by γ and γ_{TERM}, γ_{DEF}, γ_{DY}, γ_{RF}, γ_{PE}, γ_{VS}, and γ_{CP} denote the (covariance) risk prices associated with the slope of the yield curve, the corporate bond default spread, the market dividend-to-price ratio, the short-term risk-free rate, the aggregate price–earnings ratio, the value spread, and the Cochrane–Piazzesi factor. The estimates of the (covariance) risk prices are reported in the first row that corresponds to each model and the second row contains the asymptotic GMM robust t-statistics (in parentheses). The levels of statistical significance of the estimated (covariance) risk prices are indicated by italics (10% level), an underline (5% level), and boldface (1% level). The second to last column shows the average absolute pricing errors (%) and the last column presents the OLS cross-sectional R^2 value of each model. The sample period is December 1990 to December 2011.

Model	γ	γ_{TERM}	γ_{DEF}	γ_{DY}	γ_{RF}	γ_{PE}	γ_{VS}	γ_{CP}	MAE (%)	R^2_{OLS}
Panel A: SBM25										
CAPM	2.20								0.22	-0.08
	(1.42)									
HL	1.55	-161.09	-125.89						0.21	-0.03
	(0.75)	(-1.28)	(-0.70)							
P	-17.53	68.09	451.93	-7.37	-310.10				0.13	0.62
	(-1.44)	(0.39)	(1.55)	(*-1.81*)	(*-1.90*)					
CV	24.17	-80.97				-20.00	-13.21		0.10	0.76
	(0.61)	(-0.60)				(-0.53)	(-1.60)			
KLVN	3.81	0.20						-127.79	0.20	0.02
	(<u>2.13</u>)	(0.00)						(*-1.76*)		
Panel B: SM25										
CAPM	2.23								0.35	-0.28
	(1.43)									
HL	5.93	256.07	426.64						0.33	-0.14
	(<u>2.29</u>)	(*1.78*)	(1.49)							
P	29.15	-4.84	-255.47	8.99	325.07				0.17	0.68
	(<u>2.37</u>)	(-0.03)	(-1.04)	(<u>2.18</u>)	(1.55)					
CV	-42.24	-40.25				41.49	18.30		0.29	0.15
	(-1.20)	(-0.24)				(1.28)	(*1.92*)			
KLVN	0.71	-31.52						127.40	0.35	-0.25
	(0.33)	(-0.27)						(1.18)		

Table 45. Factor risk premiums for empirical risk factors (excluding excess market return).

This table reports the estimation results for the multifactor models with empirical factors using first-stage GMM. Panel A displays the results for the 25 size/book-to-market portfolios of the European stock market as test assets and Panel B presents the results for the 25 size/momentum portfolios. I do not add the market excess return to the test assets. The market (covariance) risk price is represented by γ and γ_{SMB}, γ_{HML}, γ_{UMD}, γ_L, γ_{TERM}, and γ_{DEF} denote the (covariance) risk prices associated with the size factor, the value factor, the momentum factor, the liquidity factor, the slope of the yield curve, as well as the corporate bond default spread. The estimates of the (covariance) risk prices are reported in the first row that corresponds to each model and the second row contains the asymptotic GMM robust t-statistics (in parentheses). The levels of statistical significance of the estimated (covariance) risk prices are indicated by italics (10% level), an underline (5% level), and boldface (1% level). The second to last column shows the average absolute pricing errors (%) and the last column presents the OLS cross-sectional R^2 value of each model. The sample period is December 1990 to December 2011.

Model	γ	γ_{SMB}	γ_{HML}	γ_{UMD}	γ_L	γ_{TERM}	γ_{DEF}	MAE (%)	R^2_{OLS}
Panel A: SBM25									
FF3	1.77	-2.07	7.63					0.12	0.64
	(1.12)	(-0.70)	(**3.25**)						
C	5.85	-1.12	10.53	10.64				0.09	0.80
	(2.50)	(-0.35)	(**3.81**)	(2.35)					
PS	6.45	1.54	12.11		-452.52			0.10	0.73
	(2.31)	(0.35)	(2.57)		(*-1.80*)				
FF5	5.33	3.99	9.22			-35.36	293.17	0.11	0.69
	(2.02)	(0.82)	(**2.95**)			(-0.30)	(*1.83*)		
Panel B: SM25									
FF3	3.10	3.19	-12.16					0.31	-0.10
	(*1.73*)	(0.94)	(*-1.66*)						
C	4.26	1.24	9.78	7.02				0.15	0.71
	(2.37)	(0.39)	(2.21)	(**3.48**)					
PS	7.60	9.41	-10.82		-423.84			0.32	-0.05
	(**2.62**)	(2.32)	(-1.42)		(**-2.62**)				
FF5	19.41	48.85	-9.99			-948.05	987.79	0.23	0.43
	(1.49)	(*1.83*)	(-0.67)			(-2.51)	(1.39)		

Table 46. Beta factor risk premiums for ICAPM specifications (OLS).

This table reports the estimation results of the beta factor risk premiums from OLS cross-sectional regressions. Panel A displays the results for the 25 size/book-to-market portfolios of the European stock market as test assets and Panel B presents the results for the 25 size/momentum portfolios. I add the market excess return to both sets of test assets. The estimate of RRA is denoted by γ. The beta risk price for the market factor is represented by λ_M and λ_{TERM}, λ_{DEF}, λ_{DY}, λ_{RF}, λ_{PE}, λ_{VS}, and λ_{CP} denote the beta risk prices associated with the slope of the yield curve, the corporate bond default spread, the market dividend-to-price ratio, the short-term risk-free rate, the aggregate price–earnings ratio, the value spread, and the Cochrane–Piazzesi factor. The estimates of the beta risk prices (multiplied by 100) are reported in the first row that corresponds to each model and the second row contains the Shanken (1992) t-statistics (in parentheses). The levels of statistical significance of the estimated beta risk prices are indicated by italics (10% level), an underline (5% level), and boldface (1% level). The last column presents the OLS cross-sectional R^2 value of each model. The sample period is December 1990 to December 2011.

Model	γ	λ_M	λ_{TERM}	λ_{DEF}	λ_{DY}	λ_{RF}	λ_{PE}	λ_{VS}	λ_{CP}	R^2_{OLS}
Panel A: SBM25										
CAPM	2.32	0.46								-0,07
		(1.54)								
HL	2.58	0.51	-0.04	0.01						-0,02
		(*1.74*)	(-0.54)	(0.30)						
P	2.60	0.51	0.17	0.06	-4.07	-0.18				0,50
		(*1.73*)	(*1.66*)	(<u>1.97</u>)	(<u>-2.22</u>)	(<u>-2.42</u>)				
CV	3.76	0.74	-0.07				-0.06	-5.84		0,70
		(<u>2.48</u>)	(-0.78)				(-0.16)	(**-2.89**)		
KLVN	2.51	0.50	-0.07						-0.23	0,03
		(*1.68*)	(-0.95)						(<u>-2.15</u>)	
Panel B: SM25										
CAPM	2.37	0.47								-0,27
		(1.57)								
HL	3.54	0.70	0.15	0.06						-0,08
		(<u>2.34</u>)	(1.64)	(*1.87*)						
P	3.30	0.65	-0.07	0.00	2.93	0.13				0,65
		(<u>2.20</u>)	(-1.01)	(0.14)	(*1.76*)	(<u>2.46</u>)				
CV	3.26	0.64	-0.03				0.46	4.72		0,15
		(<u>2.18</u>)	(-0.49)				(1.33)	(<u>1.99</u>)		
KLVN	2.24	0.44	0.03						0.17	-0,25
		(1.49)	(0.45)						(1.13)	

Table 47. Beta factor risk premiums for empirical risk factors (OLS).

This table reports the estimation results of the beta factor risk premiums from OLS cross-sectional regressions. Panel A displays the results for the 25 size/book-to-market portfolios of the European stock market as test assets and Panel B presents the results for the 25 size/momentum portfolios. I add the market excess return to both sets of test assets. The estimate of RRA is denoted by γ. The beta risk price for the market factor is represented by λ_M and λ_{SMB}, λ_{HML}, λ_{UMD}, λ_L, λ_{TERM}, and λ_{DEF} denote the beta risk prices associated with the size factor, the value factor, the momentum factor, the liquidity factor, the slope of the yield curve, and the corporate bond default spread. The estimates of the beta risk prices (multiplied by 100) are reported in the first row that corresponds to each model and the second row contains the Shanken (1992) t-statistics (in parentheses). The levels of statistical significance of the estimated beta risk prices are indicated by italics (10% level), an underline (5% level), and boldface (1% level). The last column presents the OLS cross-sectional R^2 value of each model. The sample period is December 1990 to December 2011.

Model	γ	λ_M	λ_{SMB}	λ_{HML}	λ_{UMD}	λ_L	λ_{TERM}	λ_{DEF}	R^2_{OLS}
Panel A: SBM25									
FF3	2.36	0.47	-0.23	0.53					0,50
		(1.60)	(-1.52)	(**2.94**)					
C	3.64	0.72	-0.20	0.53	1.84				0,74
		(<u>2.44</u>)	(-1.29)	(**2.89**)	(**3.36**)				
PS	2.71	0.54	-0.22	0.40		-0.19			0,61
		(*1.81*)	(-1.44)	(<u>2.13</u>)		(**-2.62**)			
FF5	2.97	0.59	-0.25	0.45			0.03	0.06	0,60
		(<u>2.01</u>)	(-1.64)	(<u>2.43</u>)			(0.40)	(<u>2.42</u>)	
Panel B: SM25									
FF3	2.78	0.55	0.10	-0.95					-0,08
		(*1.89*)	(0.58)	(<u>-2.29</u>)					
C	2.61	0.52	-0.08	0.44	0.82				0,64
		(*1.78*)	(-0.48)	(1.46)	(**2.99**)				
PS	2.80	0.55	0.26	-1.27		-0.16			-0,03
		(*1.89*)	(1.34)	(<u>-2.18</u>)		(<u>-2.37</u>)			
FF5	2.80	0.55	0.37	-1.41			-0.49	0.16	0,39
		(*1.76*)	(1.14)	(-1.17)			(*-1.66*)	(1.64)	

Table 48. Beta factor risk premiums for ICAPM specifications (GLS).

This table reports the estimation results of the beta factor risk premiums from GLS cross-sectional regressions. Panel A displays the results for the 25 size/book-to-market portfolios of the European stock market as test assets and Panel B presents the results for the 25 size/momentum portfolios. I add the market excess return to both sets of test assets. The estimate of RRA is denoted by γ. The beta risk price for the market factor is represented by λ_M and λ_{TERM}, λ_{DEF}, λ_{DY}, λ_{RF}, λ_{PE}, λ_{VS}, and λ_{CP} denote the beta risk prices associated with the slope of the yield curve, the corporate bond default spread, the market dividend-to-price ratio, the short-term risk-free rate, the aggregate price–earnings ratio, the value spread, and the Cochrane–Piazzesi factor. The estimates of the beta risk prices (multiplied by 100) are reported in the first row that corresponds to each model and the second row contains the Shanken (1992) t-statistics (in parentheses). The levels of statistical significance of the estimated beta risk prices are indicated by italics (10% level), an underline (5% level), and boldface (1% level). The last column presents the GLS cross-sectional R^2 value of each model. The sample period is December 1990 to December 2011.

Model	γ	λ_M	λ_{TERM}	λ_{DEF}	λ_{DY}	λ_{RF}	λ_{PE}	λ_{VS}	λ_{CP}	R^2_{GLS}
Panel A: SBM25										
CAPM	5.66	1.12								-0,06
		(**3.89**)								
HL	5.66	1.12	0.21	0.05						0,40
		(**3.89**)	(**3.17**)	(2.45)						
P	5.66	1.12	0.21	0.03	-0.07	-0.08				0,51
		(**3.89**)	(**2.92**)	(1.53)	(-0.05)	(*-1.68*)				
CV	5.66	1.12	-0.10				-0.27	-11.01		0,91
		(**3.89**)	(-1.03)				(-0.74)	(**-5.03**)		
KLVN	5.66	1.12	0.14						0.08	-0,33
		(**3.89**)	(**2.80**)						(0.92)	
Panel B: SM25										
CAPM	5.66	1.12								0,16
		(**3.89**)								
HL	5.66	1.12	0.18	0.11						0,94
		(**3.89**)	(**3.18**)	(**5.42**)						
P	5.66	1.12	0.19	0.10	5.54	-0.04				0,99
		(**3.89**)	(**3.10**)	(**4.12**)	(**3.54**)	(-0.73)				
CV	5.66	1.12	0.05				-0.29	-1.57		0,99
		(**3.89**)	(0.97)				(-0.92)	(-1.50)		
KLVN	5.66	1.12	0.11						-0.22	0,07
		(**3.89**)	(2.18)						(-1.45)	

Table 49. Beta factor risk premiums for empirical risk factors (GLS).

This table reports the estimation results of the beta factor risk premiums from GLS cross-sectional regressions. Panel A displays the results for the 25 size/book-to-market portfolios of the European stock market as test assets and Panel B presents the results for the 25 size/momentum portfolios. I add the market excess return to both sets of test assets. The estimate of RRA is denoted by γ. The beta risk price for the market factor is represented by λ_M and λ_{SMB}, λ_{HML}, λ_{UMD}, λ_L, λ_{TERM}, and λ_{DEF} denote the beta risk prices associated with the size factor, the value factor, the momentum factor, the liquidity factor, the slope of the yield curve, and the corporate bond default spread. The estimates of the beta risk prices (multiplied by 100) are reported in the first row that corresponds to each model and the second row contains the Shanken t-statistics (in parentheses). The levels of statistical significance of the estimated beta risk prices are indicated by italics (10% level), an underline (5% level), and boldface (1% level). The last column presents the GLS cross-sectional R^2 value of each model. The sample period is December 1990 to December 2011.

Model	γ	λ_M	λ_{SMB}	λ_{HML}	λ_{UMD}	λ_L	λ_{TERM}	λ_{DEF}	R^2_{GLS}
Panel A: SBM25									
FF3	5.66	1.12	-0.25	0.65					-0,13
		(**3.89**)	(*-1.67*)	(**3.70**)					
C	5.66	1.12	-0.17	0.62	3.44				0,88
		(**3.89**)	(-1.09)	(**3.45**)	(**7.34**)				
PS	5.66	1.12	-0.26	0.62		-0.07			-0,03
		(**3.89**)	(*-1.69*)	(**3.50**)		(<u>-2.25</u>)			
FF5	5.66	1.12	-0.29	0.50			0.21	0.05	0,41
		(**3.89**)	(*-1.89*)	(**2.75**)			(**2.73**)	(<u>2.31</u>)	
Panel B: SM25									
FF3	5.66	1.12	-0.30	-0.35					0,47
		(**3.89**)	(*-1.91*)	(-1.33)					
C	5.66	1.12	-0.31	-0.01	0.86				0,49
		(**3.89**)	(<u>-1.98</u>)	(-0.03)	(**3.15**)				
PS	5.66	1.12	-0.39	-0.09		0.09			0,44
		(**3.89**)	(<u>-2.43</u>)	(-0.32)		(<u>2.48</u>)			
FF5	5.66	1.12	-0.30	0.43			0.17	0.14	0,76
		(**3.89**)	(*-1.72*)	(1.01)			(<u>2.13</u>)	(**4.86**)	

Table 50. Factor risk premiums for ICAPM specifications (orthogonal factors).

This table reports the estimation results for the multifactor models with economically motivated factors using first-stage GMM when the factors are orthogonalized relative to the market factor. Panel A displays the results for the 25 size/book-to-market portfolios of the European stock market as test assets and Panel B presents the results for the 25 size/momentum portfolios. I add the market excess return to both sets of test assets. The market (covariance) risk price is represented by γ and γ_{TERM}, γ_{DEF}, γ_{DY}, γ_{RF}, γ_{PE}, γ_{VS}, and γ_{CP} denote the (covariance) risk prices associated with the slope of the yield curve, the corporate bond default spread, the market dividend-to-price ratio, the short-term risk-free rate, the aggregate price–earnings ratio, the value spread, and the Cochrane–Piazzesi factor. The estimates of the (covariance) risk prices are reported in the first row that corresponds to each model and the second row contains the asymptotic GMM robust t-statistics (in parentheses). The levels of statistical significance of the estimated (covariance) risk prices are indicated by italics (10% level), an underline (5% level), and boldface (1% level). The second to last column shows the average absolute pricing errors (%) and the last column presents the OLS cross-sectional R^2 value of each model. The sample period is December 1990 to December 2011.

Model	γ	γ_{TERM}	γ_{DEF}	γ_{DY}	γ_{RF}	γ_{PE}	γ_{VS}	γ_{CP}	MAE (%)	R^2_{OLS}
Panel A: SBM25										
HL	2.59	-63.54	50.89						0.23	-0.02
	(1.62)	(-0.54)	(0.38)							
P	2.61	132.94	594.22	-6.99	-363.24				0.15	0.50
	(1.06)	(0.68)	(*1.76*)	(*-1.71*)	(*-1.90*)					
CV	3.78	-141.60				-72.62	-16.22		0.12	0.70
	(1.46)	(-0.85)				(*-1.88*)	(-1.59)			
KLVN	2.52	-18.65						-132.29	0.22	0.03
	(1.49)	(-0.13)						(*-1.79*)		
Panel B: SM25										
HL	3.55	285.68	505.58						0.33	-0.08
	(1.59)	(*1.85*)	(1.53)							
P	3.31	45.23	-147.01	8.87	347.50				0.18	0.65
	(1.37)	(0.34)	(-0.67)	(2.16)	(*1.75*)					
CV	3.27	-62.42				-23.05	13.51		0.30	0.15
	(1.45)	(-0.47)				(-1.05)	(*1.95*)			
KLVN	2.25	-31.54						100.50	0.36	-0.25
	(1.28)	(-0.28)						(0.99)		

Table 51. Factor risk premiums for empirical risk factors (orthogonal factors).

This table reports the estimation results for the multifactor models with empirical factors using first-stage GMM when the factors are orthogonalized relative to the market factor. Panel A displays the results for the 25 size/book-to-market portfolios of the European stock market as test assets and Panel B presents the results for the 25 size/momentum portfolios. I add the market excess return to both sets of test assets. The market (covariance) risk price is represented by γ and γ_{SMB}, γ_{HML}, γ_{UMD}, γ_L, γ_{TERM}, and γ_{DEF} denote the (covariance) risk prices associated with the size factor, the value factor, the momentum factor, the liquidity factor, the slope of the yield curve, as well as the corporate bond default spread. The estimates of the (covariance) risk prices are reported in the first row that corresponds to each model and the second row contains the asymptotic GMM robust t-statistics (in parentheses). The levels of statistical significance of the estimated (covariance) risk prices are indicated by italics (10% level), an underline (5% level), and boldface (1% level). The second to last column shows the average absolute pricing errors (%) and the last column presents the OLS cross-sectional R^2 value of each model. The sample period is December 1990 to December 2011.

Model	γ	γ_{SMB}	γ_{HML}	γ_{UMD}	γ_L	γ_{TERM}	γ_{DEF}	MAE (%)	R^2_{OLS}
Panel A: SBM25									
FF3	2.37	-2.96	7.32					0.13	0.50
	(1.54)	(-1.00)	(**3.11**)						
C	3.65	-1.37	11.25	14.00				0.10	0.74
	(*1.77*)	(-0.40)	(**3.66**)	(**2.95**)					
PS	2.72	1.57	12.86		-556.52			0.12	0.61
	(1.07)	(0.31)	(2.31)		(*-1.74*)				
FF5	2.98	3.74	9.51			64.49	456.41	0.14	0.60
	(1.55)	(0.75)	(**2.91**)			(0.49)	(2.06)		
Panel B: SM25									
FF3	2.79	2.26	-13.57					0.32	-0.08
	(1.63)	(0.66)	(*-1.82*)						
C	2.62	0.27	8.13	6.97				0.16	0.64
	(1.57)	(0.09)	(*1.88*)	(**3.45**)					
PS	2.81	8.56	-12.19		-428.95			0.33	-0.03
	(1.12)	(2.06)	(-1.57)		(-2.52)				
FF5	2.81	46.85	-8.68			-842.76	1096.86	0.25	0.39
	(0.50)	(2.00)	(-0.58)			(-2.57)	(1.57)		

Table 52. Factor risk premiums for ICAPM specifications: AR(1).

This table reports the estimation results for the multifactor models with economically motivated factors using first-stage GMM when the state variables are generated by an AR(1) process. Panel A displays the results for the 25 size/book-to-market portfolios of the European stock market as test assets and Panel B presents the results for the 25 size/momentum portfolios. I add the market excess return to both sets of test assets. The market (covariance) risk price is represented by γ and γ_{TERM}, γ_{DEF}, γ_{DY}, γ_{RF}, γ_{PE}, γ_{VS}, and γ_{CP} denote the (covariance) risk prices associated with the slope of the yield curve, the corporate bond default spread, the market dividend-to-price ratio, the short-term risk-free rate, the aggregate price–earnings ratio, the value spread, and the Cochrane–Piazzesi factor. The estimates of the (covariance) risk prices are reported in the first row that corresponds to each model and the second row contains the asymptotic GMM robust t-statistics (in parentheses). The levels of statistical significance of the estimated (covariance) risk prices are indicated by italics (10% level), an underline (5% level), and boldface (1% level). The second to last column shows the average absolute pricing errors (%) and the last column presents the OLS cross-sectional R^2 value of each model. The sample period is December 1990 to December 2011.

Model	γ	γ_{TERM}	γ_{DEF}	γ_{DY}	γ_{RF}	γ_{PE}	γ_{VS}	γ_{CP}	MAE (%)	R^2_{OLS}
Panel A: SBM25										
HL	2.85	-108.26	-8.91						0.23	-0.02
	(1.38)	(-0.86)	(-0.06)							
P	-15.53	120.62	625.84	-7.30	-352.08				0.14	0.53
	(-1.30)	(0.61)	(*1.81*)	(*-1.83*)	(*-1.81*)					
CV	91.12	-181.09				-81.99	-18.46		0.13	0.68
	(2.02)	(-0.99)				(*-1.93*)	(-1.62)			
KLVN	5.51	44.73						-262.47	0.20	0.18
	(2.45)	(0.26)						(**-2.67**)		
Panel B: SM25										
HL	5.70	244.67	406.51						0.34	-0.16
	(2.28)	(*1.77*)	(1.45)							
P	31.69	22.93	-229.48	9.63	355.73				0.18	0.63
	(2.43)	(0.16)	(-0.94)	(2.19)	(*1.69*)					
CV	-51.81	-12.93				50.51	19.82		0.29	0.17
	(-1.10)	(-0.08)				(1.17)	(*1.87*)			
KLVN	0.29	-58.07						177.03	0.35	-0.22
	(0.12)	(-0.48)						(1.58)		

Table 53. Factor risk premiums for ICAPM specifications: VAR(1).

This table reports the estimation results for the multifactor models with economically motivated factors using first-stage GMM when the state variables are generated by a VAR(1) process. Panel A displays the results for the 25 size/book-to-market portfolios of the European stock market as test assets and Panel B presents the results for the 25 size/momentum portfolios. I add the market excess return to both sets of test assets. The market (covariance) risk price is represented by γ and γ_{TERM}, γ_{DEF}, γ_{DY}, γ_{RF}, γ_{PE}, γ_{VS}, and γ_{CP} denote the (covariance) risk prices associated with the slope of the yield curve, the corporate bond default spread, the market dividend-to-price ratio, the short-term risk-free rate, the aggregate price–earnings ratio, the value spread, and the Cochrane–Piazzesi factor. The estimates of the (covariance) risk prices are reported in the first row that corresponds to each model and the second row contains the asymptotic GMM robust t-statistics (in parentheses). The levels of statistical significance of the estimated (covariance) risk prices are indicated by italics (10% level), an underline (5% level), and boldface (1% level). The second to last column shows the average absolute pricing errors (%) and the last column presents the OLS cross-sectional R^2 value of each model. The sample period is December 1990 to December 2011.

Model	γ	γ_{TERM}	γ_{DEF}	γ_{DY}	γ_{RF}	γ_{PE}	γ_{VS}	γ_{CP}	MAE (%)	R^2_{OLS}
Panel A: SBM25										
HL	3.36	-88.56	74.56						0.23	-0.03
	(1.64)	(-0.65)	(0.49)							
P	-10.23	32.20	457.54	-5.46	-399.49				0.13	0.62
	(-1.25)	(0.19)	(1.45)	(*-1.85*)	(-2.23)					
CV	12.07	-201.89				-8.97	-12.32		0.13	0.55
	(0.57)	(-1.04)				(-0.43)	(*-1.85*)			
KLVN	4.13	-115.25						-238.36	0.19	0.23
	(*1.91*)	(-0.60)						(-2.53)		
Panel B: SM25										
HL	6.43	390.57	668.00						0.30	0.09
	(2.01)	(2.32)	(1.29)							
P	33.27	-178.98	-231.62	10.41	80.56				0.21	0.51
	(**2.97**)	(-1.04)	(-0.95)	(**2.68**)	(0.30)					
CV	40.10	-266.54				-37.22	15.81		0.28	0.19
	(2.02)	(-1.41)				(*-1.89*)	(2.00)			
KLVN	2.09	32.63						29.35	0.36	-0.26
	(1.26)	(0.30)						(0.35)		

References

Acharya, Viral V., and Lasse H. Pedersen, 2005, Asset pricing with liquidity risk, *Journal of Financial Economics* 77, 375–410.

Aharoni, Gil, Bruce Grundy, and Qi Zeng, 2013, Stock returns and the Miller Modigliani valuation formula: Revisiting the Fama French analysis, *Journal of Financial Economics* 110, 347–357.

Al-Horani, A., Peter F. Pope, and Andrew W. Stark, 2003, Research and development activity and expected returns in the United Kingdom, *European Finance Review* 7, 27–46.

Amihud, Yakov, 2002, Illiquidity and stock returns: Cross-section and time-series effects, *Journal of Financial Markets* 5, 31–56.

Amihud, Yakov, and Haim Mendelson, 1986, Asset pricing and the bid–ask spread, *Journal of Financial Economics* 17, 223–249.

Amihud, Yakov, Haim Mendelson, and Lasse H. Pedersen, 2005, Liquidity and asset prices, *Foundations and Trends in Finance* 1, 269–364.

Ang, Andrew, and Geert Bekaert, 2007, Stock return predictability: Is it there? *Review of Financial Studies* 20, 651–707.

Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang, 2006, The cross-section of volatility and expected returns, *Journal of Finance* 61, 259–299.

Ang, Andrew, and Jun Liu, 2004, How to discount cashflows with time-varying expected returns, *Journal of Finance* 59, 2745–2783.

Aragon, George O., and Wayne E. Ferson, 2006, Portfolio performance evaluation, *Foundations and Trends in Finance* 2, 83–190.

Artmann, Sabine, Philipp Finter, and Alexander Kempf, 2012a, Determinants of expected stock returns: Large sample evidence from the German market, *Journal of Business Finance & Accounting* 39, 758–784.

Artmann, Sabine, Philipp Finter, Alexander Kempf, Stefan Koch, and Erik Theissen, 2012b, The cross-section of German stock returns: New data and new evidence, *Schmalenbach Business Review* 64, 20–43.

Asness, Clifford S., and John M. Liew, 2014, The great divide over market efficiency. http://www.institutionalinvestor.com/Article/3315202/Asset-Management-Equities/The-Great-Divide-over-Market-Efficiency.html?ArticleId=3315202&p=1 (accessed March 18, 2014).

Asness, Clifford S., John M. Liew, and Ross L. Stevens, 1997, Parallels between the cross-sectional predictability of stock and country returns, *Journal of Portfolio Management* 23, 79–87.

Asness, Clifford S., Tobias J. Moskowitz, and Lasse H. Pedersen, 2013, Value and momentum everywhere, *Journal of Finance* 68, 929–985.

Asquith, Paul, Parag A. Pathak, and Jay R. Ritter, 2005, Short interest, institutional ownership, and stock returns, *Journal of Financial Economics* 78, 243–276.

Avramov, Doron, 2002, Stock return predictability and model uncertainty, *Journal of Financial Economics* 64, 423–458.

Avramov, Doron, Tarun Chordia, Gergana Jostova, and Alexander Philipov, 2007, Momentum and credit rating, *Journal of Finance* 62, 2503–2520.

Baker, Malcolm, and Jeffrey Wurgler, 2000, The equity share in new issues and aggregate stock returns, *Journal of Finance* 55, 2219–2257.

Baker, Malcolm, and Jeffrey Wurgler, 2006, Investor sentiment and the cross-section of stock returns, *Journal of Finance* 61, 1645–1680.

Baker, Malcolm, and Jeffrey Wurgler, 2007, Investor sentiment in the stock market, *Journal of Economic Perspectives* 21, 129–151.

Bakshi, Gurdip, and George Panayotov, 2013, Predictability of currency carry trades and asset pricing implications, *Journal of Financial Economics* 110, 139–163.

Bansal, Ravi, Varoujan Khatchatrian, and Amir Yaron, 2005, Interpretable asset markets? *European Economic Review* 49, 531–560.

Banz, Rolf W., 1981, The relationship between return and market value of common stocks, *Journal of Financial Economics* 9, 3–18.

Barber, Brad M., Terrance Odean, and Ning Zhu, 2009, Do retail trades move markets? *Review of Financial Studies* 22, 151–186.

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

Basu, Devraj, and Joëlle Miffre, 2013, Capturing the risk premium of commodity futures: The role of hedging pressure, *Journal of Banking & Finance* 37, 2652–2664.

Basu, Sanjoy, 1977, Investment performance of common stocks in relation to their price–earnings ratios: A test of the efficient market hypothesis, *Journal of Finance* 32, 663–682.

Basu, Sanjoy, 1983, The relationship between earnings' yield, market value and return for NYSE common stocks: Further evidence, *Journal of Financial Economics* 12, 129–156.

Bauer, Rob, Mathijs Cosemans, and Peter C. Schotman, 2010, Conditional asset pricing and stock market anomalies in Europe, *European Financial Management* 16, 165–190.

Beeler, Jason, and John Y. Campbell, 2012, The long-run risks model and aggregate asset prices: An empirical assessment, *Critical Finance Review* 1, 141–182.

Bekaert, Geert, and Robert J. Hodrick, 1992, Characterizing predictable components in excess returns on equity and foreign exchange markets, *Journal of Finance* 47, 467–509.

Bekaert, Geert, Robert J. Hodrick, and Xiaoyan Zhang, 2009, International stock return comovements, *Journal of Finance* 64, 2591–2626.

Belousova, Julia, and Gregor Dorfleitner, 2012, On the diversification benefits of commodities from the perspective of euro investors, *Journal of Banking & Finance* 36, 2455–2472.

Berk, Jonathan B., Richard C. Green, and Vasant Naik, 1999, Optimal investment, growth options, and security returns, *Journal of Finance* 54, 1553–1607.

Bhandari, Laxmi C., 1988, Debt/equity ratio and expected common stock returns: Empirical evidence, *Journal of Finance* 43, 507–528.

Bhojraj, Sanjeev, and Bhaskaran Swaminathan, 2006, Macromomentum: Returns predictability in international equity indices, *Journal of Business* 79, 429–451.

Black, Fischer, 1972, Capital market equilibrium with restricted borrowing, *Journal of Business* 45, 444–455.

Black, Fischer, 1993, Beta and return, *Journal of Portfolio Management* 20, 8–18.

Black, Fischer, Michael C. Jensen, and Myron Scholes, 1972, The capital asset pricing model: Some empirical tests, in Michael C. Jensen, ed.: *Studies in the Theory of Capital Markets* (Praeger Publishers, New York).

Blume, Marshall E., 1970, Portfolio theory: A step toward its practical application, *Journal of Business* 43, 152–173.

Blume, Marshall E., and Irwin Friend, 1973, A new look at the capital asset pricing model, *Journal of Finance* 28, 19–34.

Bodie, Zvi, 1976, Common stocks as a hedge against inflation, *Journal of Finance* 31, 459–470.

Bollerslev, Tim, George Tauchen, and Hao Zhou, 2009, Expected stock returns and variance risk premia, *Review of Financial Studies* 22, 4463–4492.

Bossaerts, Peter, and Pierre Hillion, 1999, Implementing statistical criteria to select return forecasting models: What do we learn? *Review of Financial Studies* 12, 405–428.

Boudoukh, Jacob, Roni Michaely, Matthew Richardson, and Michael R. Roberts, 2007a, On the importance of measuring payout yield: Implications for empirical asset pricing, *Journal of Finance* 62, 877–915.

Boudoukh, Jacob, Matthew Richardson, and Robert F. Whitelaw, 2007b, The myth of long-horizon predictability, *Review of Financial Studies* 21, 1577–1605.

Breeden, Douglas T., 1979, An intertemporal asset pricing model with stochastic consumption and investment opportunities, *Journal of Financial Economics* 7, 265–296.

Brennan, Michael J., 1958, The supply of storage, *American Economic Review* 48, 50–72.

Brennan, Michael J., Tarun Chordia, and Avanidhar Subrahmanyam, 1998, Alternative factor specifications, security characteristics, and the cross-section of expected stock returns, *Journal of Financial Economics* 49, 345–373.

Brennan, Michael J., Eduardo S. Schwartz, and Ronald Lagnado, 1997, Strategic asset allocation, *Journal of Economic Dynamics and Control* 21, 1377–1403.

Brennan, Michael J., and Avanidhar Subrahmanyam, 1996, Market microstructure and asset pricing: On the compensation for illiquidity in stock returns, *Journal of Financial Economics* 41, 441–464.

Brennan, Michael J., Ashley W. Wang, and Yihong Xia, 2004, Estimation and test of a simple model of intertemporal capital asset pricing, *Journal of Finance* 59, 1743–1775.

Callen, Jeffrey L., and Matthew R. Lyle, 2014, The term structure of implied costs of equity capital, working paper, University of Toronto, Rotman School of Management.

Campbell, John Y., 1987, Stock returns and the term structure, *Journal of Financial Economics* 18, 373–399.

Campbell, John Y., 1991, A variance decomposition for stock returns, *Economic Journal* 101, 157–179.

Campbell, John Y., 1993, Intertemporal asset pricing without consumption data, *American Economic Review* 83, 487–512.

Campbell, John Y., 1996, Understanding risk and return, *Journal of Political Economy* 104, 298–345.

Campbell, John Y., 2000, Asset pricing at the millennium, *Journal of Finance* 55, 1515–1567.

Campbell, John Y., 2014, Empirical asset pricing: Eugene Fama, Lars Peter Hansen, and Robert Shiller, *Scandinavian Journal of Economics* 116, 593–634.

Campbell, Sean D., Morris A. Davis, Joshua Gallin, and Robert F. Martin, 2009, What moves housing markets: A variance decomposition of the rent–price ratio, *Journal of Urban Economics* 66, 90–102.

Campbell, John Y., Stefano Giglio, and Christopher Polk, 2013, Hard times, *Review of Asset Pricing Studies* 3, 95–132.

Campbell, John Y., Jens Hilscher, and Jan Szilagyi, 2008, In search of distress risk, *Journal of Finance* 63, 2899–2939.

Campbell, John Y., Andrew W. Lo, and Archie C. MacKinlay, 1997, *The Econometrics of Financial Markets* (Princeton University Press, Princeton, NJ).

Campbell, John Y., and Robert J. Shiller, 1988b, Stock prices, earnings, and expected dividends, *Journal of Finance* 43, 661–676.

Campbell, John Y., and Robert J. Shiller, 1988a, The dividend–price ratio and expectations of future dividends and discount factors, *Review of Financial Studies* 1, 195–228.

Campbell, John Y., and Robert J. Shiller, 1991, Yield spreads and interest rate movements: A bird's eye view, *Review of Economic Studies* 58, 495–514.

Campbell, John Y., and Samuel B. Thompson, 2008, Predicting excess stock returns out of sample: Can anything beat the historical average? *Review of Financial Studies* 21, 1509–1531.

Campbell, John Y., and Luis M. Viceira, 1999, Consumption and portfolio decisions when expected returns are time varying, *Quarterly Journal of Economics* 114, 433–495.

Campbell, John Y., and Tuomo Vuolteenaho, 2004, Bad beta, good beta, *American Economic Review* 94, 1249–1275.

Carhart, Mark M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57–82.

Chan, K. C., and Nai-Fu Chen, 1991, Structural and return characteristics of small and large firms, *Journal of Finance* 46, 1467–1484.

Chan, K. C., Yasushi Hamao, and Josef Lakonishok, 1991, Fundamentals and stock returns in Japan, *Journal of Finance* 46, 1739–1764.Chen, Long, and Xinlei Zhao, 2009, Return decomposition, *Review of Financial Studies* 22, 5213–5249.

Chen, Nai-Fu, Richard Roll, and Stephen A. Ross, 1986, Economic forces and the stock market, *Journal of Business* 59, 383–403.

Chordia, Tarun, and Lakshmanan Shivakumar, 2002, Momentum, business cycle, and time-varying expected returns, *Journal of Finance* 57, 985–1019.

Chordia, Tarun, Avanidhar Subrahmanyam, and V. R. Anshuman, 2001, Trading activity and expected stock returns, *Journal of Financial Economics* 59, 3–32.

Chui, Andy C. W., Sheridan Titman, and K. C. John Wei, 2010, Individualism and momentum around the world, *Journal of Finance* 65, 361–392.

Clark, Todd E., and Kenneth D. West, 2007, Approximately normal tests for equal predictive accuracy in nested models, *Journal of Econometrics* 138, 291–311.

Cochrane, John H., 1991, Production-based asset pricing and the link between stock returns and economic fluctuations, *Journal of Finance* 46, 209–237.

Cochrane, John H., 2005, *Asset Pricing* (Princeton University Press, Princeton, NJ).

Cochrane, John H., 2008, The dog that did not bark: A defense of return predictability, *Review of Financial Studies* 21, 1533–1575.

Cochrane, John H., 2011, Presidential address: Discount rates, *Journal of Finance* 66, 1047–1108.

Cochrane, John H., 2014, A mean–variance benchmark for intertemporal portfolio theory, *Journal of Finance* 69, 1–49.

Cochrane, John H., and Monika Piazzesi, 2005, Bond risk premia, *American Economic Review* 95, 138–160.

Cohen, Lauren, and Andrea Frazzini, 2008, Economic links and predictable returns, *Journal of Finance* 63, 1977–2011.

Conrad, Jennifer, and Gautam Kaul, 1988, Time-variation in expected returns, *Journal of Business* 61, 409–425.

Constantinides, George M., 1986, Capital market equilibrium with transaction costs, *Journal of Political Economy* 94, 842–862.

Cooper, Michael J., Huseyin Gulen, and Michael J. Schill, 2008, Asset growth and the cross-section of stock returns, *Journal of Finance* 63, 1609–1651.

Cooper, Michael J., Roberto C. Gutierrez, Jr., and Allaudeen Hameed, 2004, Market states and momentum, *Journal of Finance* 59, 1345–1365.

Cootner, Paul H., 1960, Returns to speculators: Telser versus Keynes, *Journal of Political Economy* 68, 396–404.

Coppola, Andrea, 2008, Forecasting oil price movements: Exploiting the information in the futures market, *Journal of Futures Markets* 28, 34–56.

Coval, Joshua D., and Tobias J. Moskowitz, 2001, The geography of investment: Informed trading and asset prices, *Journal of Political Economy* 109, 811–841.

Da, Zhi, Re-Jin Guo, and Ravi Jagannathan, 2012, CAPM for estimating the cost of equity capital: Interpreting the empirical evidence, *Journal of Financial Economics* 103, 204–220.

Daniel, Kent, and Sheridan Titman, 1997, Evidence on the characteristics of cross sectional variation in stock returns, *Journal of Finance* 52, 1–33.

Daniel, Kent, and Sheridan Titman, 2006, Market reactions to tangible and intangible information, *Journal of Finance* 61, 1605–1643.

Daniel, Kent, and Sheridan Titman, 2012, Testing factor-model explanations of market anomalies, *Critical Finance Review* 1, 103–139.

Datar, Vinay T., Y. Naik, Narayan, and Robert Radcliffe, 1998, Liquidity and stock returns: An alternative test, *Journal of Financial Markets* 1, 203–219.

DeBondt, Werner F. M., and Richard H. Thaler, 1985, Does the stock market overreact? *Journal of Finance* 40, 793–805.

DeBondt, Werner F. M., and Richard H. Thaler, 1987, Further evidence on investor overreaction and stock market seasonality, *Journal of Finance* 42, 557–581.

De Roon, Frans A., Theo E. Nijman, and Chris Veld, 2000, Hedging pressure effects in futures markets, *Journal of Finance* 55, 1437–1456.

Dichev, Ilia D., 1998, Is the risk of bankruptcy a systematic risk? *Journal of Finance* 53, 1131–1147.

Douglas, George W., 1968, *Risk in the Equity Markets: An Empirical Appraisal of Market Efficiency* (University Microfilms, Ann Arbor, MI).

Drechsler, Itamar, and Amir Yaron, 2010, What's vol got to do with it, *Review of Financial Studies* 24, 1–45.

Dusak, Katherine, 1973, Futures trading and investor returns: An investigation of commodity market risk premiums, *Journal of Political Economy* 81, 1387–1406.

Easley, David, Soeren Hvidkjaer, and Maureen O'Hara, 2002, Is information risk a determinant of asset returns? *Journal of Finance* 57, 2185–2221.

Economist, The, 2010, Economics focus: Data birth, http://www.economist.com/node/17519706 (accessed May 2, 2014).

Epstein, Larry G., and Stanley E. Zin, 1989, Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework, *Econometrica* 57, 937–969.

Epstein, Larry G., and Stanley E. Zin, 1991, Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis, *Journal of Political Economy* 99, 263–286.

Erb, Claude B., and Campbell R. Harvey, 2006, The strategic and tactical value of commodity futures, *Financial Analysts Journal* 62, 69–97.

Fama, Eugene F., 1965, The behavior of stock-market prices, *Journal of Business* 38, 34–105.

Fama, Eugene F., 1970a, Efficient capital markets: A review of theory and empirical work, *Journal of Finance* 25, 383–417.

Fama, Eugene F., 1970b, Multiperiod consumption-investment decisions, *American Economic Review* 60, 163–174.

Fama, Eugene F., 1981, Stock returns, real activity, inflation, and money, *American Economic Review* 71, 545–565.

Fama, Eugene F., 1984, Forward and spot exchange rates, *Journal of Monetary Economics* 14, 319–338.

Fama, Eugene F., 1986, Term premiums and default premiums in money markets, *Journal of Financial Economics* 17, 175–196.

Fama, Eugene F., 1991, Efficient capital markets: II, *Journal of Finance* 46, 1575–1617.

Fama, Eugene F., 1996, Multifactor portfolio efficiency and multifactor asset pricing, *Journal of Financial and Quantitative Analysis* 31, 441–465.

Fama, Eugene F., and Robert R. Bliss, 1987, The information in long-maturity forward rates, *American Economic Review* 77, 680–692.

Fama, Eugene F., and Marshall E. Blume, 1966, Filter rules and stock-market trading, *Journal of Business* 39, 226–241.

Fama, Eugene F., and Kenneth R. French, 1987, Commodity futures prices: Some evidence on forecast power, premiums, and the theory of storage, *Journal of Business* 60, 55–73.

Fama, Eugene F., and Kenneth R. French, 1988b, Dividend yields and expected stock returns, *Journal of Financial Economics* 22, 3–25.

Fama, Eugene F., and Kenneth R. French, 1988a, Permanent and temporary components of stock prices, *Journal of Political Economy* 96, 246–273.

Fama, Eugene F., and Kenneth R. French, 1989, Business conditions and expected returns on stocks and bonds, *Journal of Financial Economics* 25, 23–49.

Fama, Eugene F., and Kenneth R. French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427–465.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F., and Kenneth R. French, 1995, Size and book-to-market factors in earnings and returns, *Journal of Finance* 50, 131–155.

Fama, Eugene F., and Kenneth R. French, 1996, Multifactor explanations of asset pricing anomalies, *Journal of Finance* 51, 55–84.

Fama, Eugene F., and Kenneth R. French, 1997, Industry costs of equity, *Journal of Financial Economics* 43, 153–193.

Fama, Eugene F., and Kenneth R. French, 1998, Value versus growth: The international evidence, *Journal of Finance* 53, 1975–1999.

Fama, Eugene F., and Kenneth R. French, 2004, The capital asset pricing model: Theory and evidence, *Journal of Economic Perspectives* 18, 25–46.

Fama, Eugene F., and Kenneth R. French, 2010, Luck versus skill in the cross-section of mutual fund returns, *Journal of Finance* 65, 1915–1947.

Fama, Eugene F., and Kenneth R. French, 2012, Size, value, and momentum in international stock returns, *Journal of Financial Economics* 105, 457–472.

Fama, Eugene F., and Kenneth R. French, 2013, A five-factor asset pricing model, working paper, University of Chicago, Dartmouth College.

Fama, Eugene F., and Robert Litterman, 2012, An experienced view on markets and investing, *Financial Analysts Journal* 68, 15–19.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607–636.

Fama, Eugene F., and G. W. Schwert, 1977, Asset returns and inflation, *Journal of Financial Economics* 5, 115–146.

Fang, Lily, and Joel Peress, 2009, Media coverage and the cross-section of stock returns, *Journal of Finance* 64, 2023–2052.

Favero, Carlo A., Arie E. Gozluklu, and Andrea Tamoni, 2011, Demographic trends, the dividend-price ratio, and the predictability of long-run stock market returns, *Journal of Financial & Quantitative Analysis* 46, 1493–1520.

Ferreira, Miguel A., and Pedro Santa-Clara, 2011, Forecasting stock market returns: The sum of the parts is more than the whole, *Journal of Financial Economics* 100, 514–537.

Ferson, Wayne E., and Campbell R. Harvey, 1991, The variation of economic risk premiums, *Journal of Political Economy* 99, 385–415.

Ferson, Wayne E., and Campbell R. Harvey, 1993, The risk and predictability of international equity returns, *Review of Financial Studies* 6, 527–566.

Ferson, Wayne E., Sergei Sarkissian, and Timothy Simin, 1999, The alpha factor asset pricing model: A parable, *Journal of Financial Markets* 2, 49–68.

Frankel, Richard, and Charles M. C. Lee, 1998, Accounting valuation, market expectation, and cross-sectional stock returns, *Journal of Accounting and Economics* 25, 283–319.

Friend, Irwin, and Marshall E. Blume, 1970, Measurement of portfolio performance under uncertainty, *American Economic Review* 60, 561–575.

Fu, Fangjian, 2009, Idiosyncratic risk and the cross-section of expected stock returns, *Journal of Financial Economics* 91, 24–37.

Gebhardt, William R., Charles M. C. Lee, and Bhaskaran Swaminathan, 2001, Toward an implied cost of capital, *Journal of Accounting Research* 39, 135–176.

Gerard, Bruno, and Guojun Wu, 2006, How important is intertemporal risk for asset allocation? *Journal of Business* 79, 2203–2241.

Gibbons, Michael R., Stephen A. Ross, and Jay Shanken, 1989, A test of the efficiency of a given portfolio, *Econometrica* 57, 1121–1152.

Gomes, Joao, Leonid Kogan, and Lu Zhang, 2003, Equilibrium cross section of returns, *Journal of Political Economy* 111, 693–732.

Gompers, Paul, Joy Ishii, and Andrew Metrick, 2003, Corporate governance and equity prices, *Quarterly Journal of Economics* 118, 107–156.

Gonedes, Nicholas J., 1973, Evidence on the information content of accounting numbers: accounting-based and market-based estimates of systematic risk, *Journal of Financial & Quantitative Analysis* 8, 407–443.

Gorton, Gary B., Fumio Hayashi, and K. G. Rouwenhorst, 2013, The fundamentals of commodity futures returns, *Review of Finance* 17, 35–105.

Gourinchas, Pierre-Olivier, and Hélène Rey, 2007, International financial adjustment, *Journal of Political Economy* 115, 665–703.

Goyal, Amit, 2012, Empirical cross-sectional asset pricing: A survey, *Financial Markets and Portfolio Management* 26, 3–38.

Gregory, Alan, and Maria Michou, 2009, Industry cost of equity capital: UK evidence, *Journal of Business Finance & Accounting* 36, 679–704.

Gregory, Alan, Rajesh Tharyan, and Angela Christidis, 2013, Constructing and testing alternative versions of the Fama–French and Carhart models in the UK, *Journal of Business Finance & Accounting* 40, 172–214.

Griffin, John M., 2002, Are the Fama and French Factors global or country specific? *Review of Financial Studies* 15, 783–803.

Griffin, John M., Xiuqing Ji, and J. S. Martin, 2003, Momentum investing and business cycle risk: Evidence from pole to pole, *Journal of Finance* 58, 2515–2547.

Grinblatt, Mark, and Tobias J. Moskowitz, 2004, Predicting stock price movements from past returns: The role of consistency and tax-loss selling, *Journal of Financial Economics* 71, 541–579.

Guo, Hui, 2006a, On the out-of-sample predictability of stock market returns, *Journal of Business* 79, 645–670.

Guo, Hui, 2006b, Time-varying risk premia and the cross section of stock returns, *Journal of Banking & Finance* 30, 2087–2107.

Guo, Hui, and Robert Savickas, 2008, Average idiosyncratic volatility in G7 Countries, *Review of Financial Studies* 21, 1259–1296.

Hahn, Jaehoon, and Hangyong Lee, 2006, Yield spreads as alternative risk factors for size and book-to-market, *Journal of Financial & Quantitative Analysis* 41, 245–269.

Hansen, Lars P., 1982, Large sample properties of generalized method of moments estimators, *Econometrica* 50, 1029–1054.

Hansen, Lars P., and Robert J. Hodrick, 1980, Forward exchange rates as optimal predictors of future spot rates: an econometric analysis, *Journal of Political Economy* 88, 829–853.

Hansen, Lars P., and Scott F. Richard, 1987, The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing Models, *Econometrica* 55, 587–613.

Harvey, Campbell R., 1995, Predictable risk and returns in emerging markets, *Review of Financial Studies* 8, 773–816.

Haugen, Robert A., 1995, *The New Finance: The Case against Efficient Markets* (Prentice Hall, Englewood Cliffs, NJ).

Haugen, Robert A., and Nardin L. Baker, 1996, Commonality in the determinants of expected stock returns, *Journal of Financial Economics* 41, 401–439.

Heaton, John, and Deborah J. Lucas, 1996, Evaluating the effects of incomplete markets on risk sharing and asset pricing, *Journal of Political Economy* 104, 443–487.

Henriksson, Roy D., and Robert C. Merton, 1981, On market timing and investment performance. II. Statistical procedures for evaluating forecasting skills, *Journal of Business* 54, 513–533.

Hicks, John R., 1939, *Value and Capital* (Oxford University Press, Cambridge, UK).

Hjalmarsson, Erik, 2010, Predicting global stock returns, *Journal of Financial & Quantitative Analysis* 45, 49–80.

Hjalmarsson, Erik, 2011, New methods for inference in long-horizon regressions, *Journal of Financial & Quantitative Analysis* 46, 815–839.

Hong, Harrison, Terence Lim, and Jeremy C. Stein, 2000, Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies, *Journal of Finance* 55, 265–295.

Hong, Harrison, and Jeremy C. Stein, 1999, A unified theory of underreaction, momentum trading, and overreaction in asset markets, *Journal of Finance* 54, 2143–2184.

Hong, Harrison, and Motohiro Yogo, 2012, What does futures market interest tell us about the macroeconomy and asset prices? *Journal of Financial Economics* 105, 473–490.

Hou, Kewei, G. A. Karolyi, and Bong-Chan Kho, 2011, What factors drive global stock returns? *Review of Financial Studies* 24, 2527–2574.

Hou, Kewei, and Tobias J. Moskowitz, 2005, Market frictions, price delay, and the cross-section of expected returns, *Review of Financial Studies* 18, 981–1020.

Hou, Kewei, and David T. Robinson, 2006, Industry concentration and average stock returns, *Journal of Finance* 61, 1927–1956.

Huang, Ming, 2003, Liquidity shocks and equilibrium liquidity premia, *Journal of Economic Theory* 109, 104–129.

Huberman, Gur, and Shmuel Kandel, 1987, Mean–variance spanning, *Journal of Finance* 42, 873–888.

Hvidkjaer, Soeren, 2006, A trade-based analysis of momentum, *Review of Financial Studies* 19, 457–491.

Hvidkjaer, Soeren, 2008, Small trades and the cross-section of stock returns, *Review of Financial Studies* 21, 1123–1151.

Ince, Ozgur S., and Porter, R. Burt, 2006, Individual equity return data from Thomson Datastream: Handle with care! *Journal of Financial Research* 29, 463–479.

Jacquier, Eric, and Cédric Okou, 2014, Disentangling continuous volatility from jumps in long-run risk-return relationships, *Journal of Financial Econometrics* 12, 544–583.

Jagannathan, Ravi, Ernst Schaumburg, and Guofu Zhou, 2010a, Cross-sectional asset pricing tests, *Annual Review of Financial Economics* 2, 49–74.

Jagannathan, Ravi, Georgios Skoulakis, and Zhenyu Wang, 2010b, The analysis of the cross-section of security returns, in Yacine Aït-Sahalia, and Lars Peter Hansen, eds.: *Handbook of Financial Econometrics Tools and Techniques* (North-Holland/Elsevier, Amsterdam).

Jagannathan, Ravi, and Zhenyu Wang, 1996, The conditional CAPM and the cross-section of expected returns, *Journal of Finance* 51, 3–53.

Jegadeesh, Narasimhan, 1990, Evidence of predictable behavior of security returns, *Journal of Finance* 45, 881–898.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65–91.

Jensen, Gerald R., Robert R. Johnson, and Jeffrey M. Mercer, 2000, Efficient use of commodity futures in diversified portfolios, *Journal of Futures Markets* 20, 489–506.

Jensen, Gerald R., Robert R. Johnson, and Jeffrey M. Mercer, 2002, Tactical asset allocation and commodity futures, *Journal of Portfolio Management* 28, 100–111.

Jensen, Michael C., 1968, The performance of mutual funds in the period 1945–1964, *Journal of Finance* 23, 389–416.

Johannes, Michael, Arthur Korteweg, and Nicholas Polson, 2014, Sequential learning, predictability, and optimal portfolio returns, *Journal of Finance* 69, 611–644.

Johnson, Shane A., Theodore C. Moorman, and Sorin Sorescu, 2009, A reexamination of corporate governance and equity prices, *Review of Financial Studies* 22, 4753–4786.

Jones, Charles M., and Owen A. Lamont, 2002, Short-sale constraints and stock returns, *Limits on Arbitrage* 66, 207–239.

Kaldor, Nicholas, 1939, Speculation and economic stability, *Review of Economic Studies* 7, 1–27.

Karolyi, G. A., and René M. Stulz, 2003, Are assets priced globally or locally? in George M. Constantinides, M. Harris, and René M. Stulz, eds.: *Handbook of the Economics of Finance* (North-Holland/Elsevier, Amsterdam).

Keim, Donald B., and Robert F. Stambaugh, 1986, Predicting returns in the stock and bond markets, *Journal of Financial Economics* 17, 357–390.

Kelly, Bryan, and Seth Pruitt, 2013, Market expectations in the cross-section of present values, *Journal of Finance* 68, 1721–1756.

Keynes, John M., 1930, *A Treatise on Money* (Macmillan, London).

Kho, Bong-Chan, 1996, Time-varying risk premia, volatility, and technical trading rule profits: Evidence from foreign currency futures markets, *Journal of Financial Economics* 41, 249–290.

Koijen, Ralph S., Hanno N. Lustig, and Stijn G. van Nieuwerburgh, 2010, The cross-section and time-series of stock and bond returns, working paper, University of Chicago, University of California at Los Angeles, New York University.

Koijen, Ralph S., and Stijn G. van Nieuwerburgh, 2011, Predictability of returns and cash flows, *Annual Review of Financial Economics* 3, 467–491.

Kosowski, Robert, Allan Timmermann, Russ Wermers, and Hal White, 2006, Can mutual fund "stars" really pick stocks? New evidence from a bootstrap analysis, *Journal of Finance* 61, 2551–2595.

Kothari, S. P., and Jay Shanken, 1997, Book-to-market, dividend yield, and expected market returns: a time-series analysis, *Journal of Financial Economics* 44, 169–203.

Kothari, S. P., Jay Shanken, and Richard G. Sloan, 1995, Another look at the cross-section of expected stock returns, *Journal of Finance* 50, 185–224.

Lakonishok, Josef, and Alan C. Shapiro, 1986, Systematic risk, total risk and size as determinants of stock market returns, *Journal of Banking & Finance* 10, 115–132.

Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny, 1994, Contrarian investment, extrapolation, and risk, *Journal of Finance* 49, 1541–1578.

Lamont, Owen A., 1998, Earnings and expected returns, *Journal of Finance* 53, 1563–1587.

La Porta, Rafael, 1996, Expectations and the cross-section of stock returns, *Journal of Finance* 51, 1715–1742.

LeBaron, Blake, 1999, Technical trading rule profitability and foreign exchange intervention, *Journal of International Economics* 49, 125–143.

Lee, Kuan-Hui, 2011, The world price of liquidity risk, *Journal of Financial Economics* 99, 136–161.

Lesmond, David A., Joseph P. Ogden, and Charles A. Trzcinka, 1999, A new estimate of transaction costs, *Review of Financial Studies* 12, 1113–1141.

Lettau, Martin, and Sydney C. Ludvigson, 2001a, Consumption, aggregate wealth, and expected stock returns, *Journal of Finance* 56, 815–849.

Lettau, Martin, and Sydney C. Ludvigson, 2001b, Resurrecting the (C)CAPM: A cross-sectional test when risk premia are time-varying, *Journal of Political Economy* 109, 1238–1287.

Lettau, Martin, and Sydney C. Ludvigson, 2010, Measuring and modeling variation in the risk-return trade-off, in Yacine Aït-Sahalia, and Lars Peter Hansen, eds.: *Handbook of Financial Econometrics Tools and Techniques* (North-Holland/Elsevier, Amsterdam).

Lewellen, Jonathan, 1999, The time-series relations among expected return, risk, and book-to-market, *Journal of Financial Economics* 54, 5–43.

Lewellen, Jonathan, Stefan Nagel, and Jay Shanken, 2010, A skeptical appraisal of asset pricing tests, *Journal of Financial Economics* 96, 175–194.

Lewis, Karen K., 2011, Global asset pricing, *Annual Review of Financial Economics* 3, 435–466.

Li, Bo, Qian Sun, and Changyun Wang, 2014, Liquidity, liquidity risk and stock returns: Evidence from Japan, *European Financial Management* 20, 126–151.

Li, Yan, David T. Ng, and Bhaskaran Swaminathan, 2013, Predicting market returns using aggregate implied cost of capital, *Journal of Financial Economics* 110, 419–436.

Liew, Jimmy, and Maria Vassalou, 2000, Can book-to-market, size and momentum be risk factors that predict economic growth? *Journal of Financial Economics* 57, 221–245.

Lintner, John, 1965a, The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, *Review of Economics & Statistics* 47, 13–37.

Lintner, John, 1965b, Security prices, risk, and maximal gains from diversification, *Journal of Finance* 20, 587–615.

Liu, Weimin, 2006, A liquidity-augmented capital asset pricing model, *Journal of Financial Economics* 82, 631–671.

Lo, Andrew W., and Archie C. MacKinlay, 1988, Stock market prices do not follow random walks: Evidence from a simple specification test, *Review of Financial Studies* 1, 41–66.

Lo, Andrew W., and Archie C. MacKinlay, 1990, Data-snooping biases in tests of financial asset pricing models, *Review of Financial Studies* 3, 431–467.

Lo, Andrew W., Harry Mamaysky, and Jiang Wang, 2004, Asset prices and trading volume under fixed transactions costs, *Journal of Political Economy* 112, 1054–1090.

Lo, Andrew W., and Jiang Wang, 2006, Trading volume: Implications of an intertemporal capital asset pricing model, *Journal of Finance* 61, 2805–2840.

Long, John B., 1974, Stock prices, inflation, and the term structure of interest rates, *Journal of Financial Economics* 1, 131–170.

Lustig, Hanno N., and Stijn G. van Nieuwerburgh, 2005, Housing collateral, consumption insurance, and risk premia: An empirical perspective, *Journal of Finance* 60, 1167–1219.

Lustig, Hanno N., Nikolai Roussanov, and Adrien Verdelhan, 2014, Countercyclical currency risk premia, *Journal of Financial Economics* 111, 527–553.

Lutzenberger, Fabian T., 2014a, The predictability of aggregate returns on commodity futures, *Review of Financial Economics* 23, 120-130.

Lutzenberger, Fabian T., 2014b, Multifactor models and their consistency with the ICAPM: Evidence from the European stock market, Discussion Paper WI-430, University of Augsburg, *European Financial Management,* forthcoming.

Lutzenberger, Fabian T., 2014c, Industry cost of equity capital: European evidence for multifactor models, Discussion Paper WI-434, University of Augsburg.

MacKinlay, Archie C., 1995, Multifactor models do not explain deviations from the CAPM, *Journal of Financial Economics* 38, 3–28.

Maio, Paulo, 2013, Return decomposition and the Intertemporal CAPM, *Journal of Banking & Finance* 37, 4958–4972.

Maio, Paulo, and Pedro Santa-Clara, 2012, Multifactor models and their consistency with the ICAPM, *Journal of Financial Economics* 106, 586–613.

Markowitz, Harry M., 1952, Portfolio selection, *Journal of Finance* 7, 77–91.

Markowitz, Harry M., 1959, *Portfolio Selection: Efficient Diversification of Investments* (John Wiley & Sons, New York).

Mehra, Rajnish, and Edward C. Prescott, 1985, The equity premium: a puzzle, *Journal of Monetary Economics* 15, 145–161.

Menzly, Lior, Tano Santos, and Pietro Veronesi, 2004, Understanding predictability, *Journal of Political Economy* 112, 1–47.

Merton, Robert C., 1971, Optimum consumption and portfolio rules in a continuous-time model, *Journal of Economic Theory* 3, 373–413.

Merton, Robert C., 1973, An intertemporal capital asset pricing model, *Econometrica* 41, 867–887.

Miffre, Joëlle, and Georgios Rallis, 2007, Momentum strategies in commodity futures markets, *Journal of Banking & Finance* 31, 1863–1886.

Miller, Merton, and Myron Scholes, 1972, Rates of return in relation to risk: A reexamination of some recent findings, in Michael C. Jensen, ed.: *Studies in the Theory of Capital Markets* (Praeger Publishers, New York).

Moskowitz, Tobias J., and Mark Grinblatt, 1999, Do industries explain momentum? *Journal of Finance* 54, 1249–1290.

Moskowitz, Tobias J., Yao H. Ooi, and Lasse H. Pedersen, 2012, Time series momentum, *Journal of Financial Economics* 104, 228–250.

Narayan, Paresh K., Seema Narayan, and Susan S. Sharma, 2013, An analysis of commodity markets: What gain for investors? *Journal of Banking & Finance* 37, 3878–3889.

Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

Novy-Marx, Robert, 2013, The other side of value: The gross profitability premium, *Journal of Financial Economics* 108, 1–28.

Novy-Marx, Robert, 2014, Predicting anomaly performance with politics, the weather, global warming, sunspots, and the stars, *Journal of Financial Economics* 112, 137–146.

Ozoguz, Arzu, 2009, Good times or bad times? Investors' uncertainty and stock returns, *Review of Financial Studies* 22, 4377–4422.

Pástor, Lubos, 2000, Portfolio selection and asset pricing models, *Journal of Finance* 55, 179–223.

Pástor, Lubos, and Robert F. Stambaugh, 2003, Liquidity risk and expected stock returns, *Journal of Political Economy* 111, 642–685.

Patton, Andrew J., and Allan Timmermann, 2010, Monotonicity in asset returns: New tests with applications to the term structure, the CAPM, and portfolio sorts, *Journal of Financial Economics* 98, 605–625.

Paye, Bradley S., and Allan Timmermann, 2006, Instability of return prediction models, *Journal of Empirical Finance* 13, 274–315.

Pennacchi, George G., 2008, *Theory of Asset Pricing* (Pearson/Addison-Wesley, Boston).

Perez-Quiros, Gabriel, and Allan Timmermann, 2000, Firm size and cyclical variations in stock returns, *Journal of Finance* 55, 1229–1262.

Pesaran, Mohammad H., and Allan Timmermann, 1992, Is value riskier than growth? *Journal of Business & Economic Statistics* 10, 461–465.

Petkova, Ralitsa, 2006, Do the Fama–French factors proxy for innovations in predictive variables? *Journal of Finance* 61, 581–612.

Piazzesi, Monika, Martin Schneider, and Selale Tuzel, 2007, Housing, consumption and asset pricing, *Journal of Financial Economics* 83, 531–569.

Piazzesi, Monika, and Eric T. Swanson, 2008, Futures prices as risk-adjusted forecasts of monetary policy, *Journal of Monetary Economics* 55, 677–691.

Plazzi, Alberto, Walter Torous, and Rossen Valkanov, 2010, Expected returns and expected growth in rents of commercial real estate, *Review of Financial Studies* 23, 3469–3519.

Polk, Christopher, Samuel Thompson, and Tuomo Vuolteenaho, 2006, Cross-sectional forecasts of the equity premium, *Journal of Financial Economics* 81, 101–141.

Pontiff, Jeffrey, and Lawrence D. Schall, 1998, Book-to-market ratios as predictors of market returns, *Journal of Financial Economics* 49, 141–160.

Poterba, James M., and Lawrence H. Summers, 1988, Mean reversion in stock prices: Evidence and implications, *Journal of Financial Economics* 22, 27–59.

Pratt, Shannon P., and Roger J. Grabowski, 2010, *Cost of Capital. Application and Examples* (Wiley, Hoboken, NJ).

Prokopczuk, Marcel, and Lazaros Symeonidis, 2013, The economic drivers of time-varying commodity market volatility, working paper, Zeppelin University, University of Reading.

Rangvid, Jesper, 2006, Output and expected returns, *Journal of Financial Economics* 81, 595–624.

Rapach, David E., Jack K. Strauss, and Guofu Zhou, 2010, Out-of-sample equity premium prediction: Combination forecasts and links to the real economy, *Review of Financial Studies* 23, 821–862.

Reinganum, Marc R., 1981, A new empirical perspective on the CAPM, *Journal of Financial and Quantitative Analysis* 16, 439.

Roll, Richard, 1977, A critique of the asset pricing theory's tests Part I: On past and potential testability of the theory, *Journal of Financial Economics* 4, 129–176.

Rosenberg, Barr, Kenneth Reid, and Ronald Lanstein, 1985, Persuasive evidence of market inefficiency, *Journal of Portfolio Management* 11, 9–17.

Ross, Stephen A., 1976, The arbitrage theory of capital asset pricing, *Journal of Economic Theory* 13, 341–360.

Rouwenhorst, K. G., 1998, International momentum strategies, *Journal of Finance* 53, 267–284.

Royal Swedish Academy of Sciences, 2013, Understanding asset prices: Scientific background on the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2013 compiled by the Economic Sciences Prize Committee of the Royal Swedish Academy of Sciences, http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/2013/advanced-economicsciences2013.pdf (accessed July 31, 2014).

Rozeff, Michael S., 1984, Dividend yields are equity risk premiums, *Journal of Portfolio Management* 11, 68–75.

Rubinstein, Mark E., 1976, The strong case for the generalized logarithmic utility model as the premier model of financial markets, *Journal of Finance* 31, 551–571.

Santos, Tano, and Pietro Veronesi, 2005, Labor income and predictable stock returns, *Review of Financial Studies* 19, 1–44.

Schrimpf, Andreas, Michael Schröder, and Richard Stehle, 2007, Cross-sectional tests of conditional asset pricing models: Evidence from the German stock market, *European Financial Management* 13, 880–907.

Shanken, Jay, 1990, Intertemporal asset pricing: An empirical investigation, *Journal of Econometrics* 45, 99–120.

Shanken, Jay, 1992, On the estimation of beta-pricing models, *Review of Financial Studies* 5, 1–55.

Sharpe, William F., 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* 19, 425–442.

Shiller, Robert J., 1984, Stock prices and social dynamics, *Brookings Papers on Economic Activity* 2, 457–510.

Sloan, Richard G., 1996, Do stock prices fully reflect information in accruals and cash flows about future earnings? *Accounting Review* 71, 289–315.

Stambaugh, Robert F., 1982, On the exclusion of assets from tests of the two-parameter model: A sensitivity analysis, *Journal of Financial Economics* 10, 237–268.

Stattman, Dennis, 1980, Book values and stock returns, *The Chicago MBA: A Journal of Selected Papers* 4, 25–45.

Su, Xiaojing, Tao Wang, and Jian Yang, 2009, Out-of-sample predictability in international equity markets: A model selection approach, *Financial Review* 44, 559–582.

Subrahmanyam, Avanidhar, 2010, The cross-section of expected stock returns: What have we learnt from the past twenty-five years of research? *European Financial Management* 16, 27–42.

Subrahmanyam, Avanidhar, 2013, Comments and perspectives on ‘The Capital Asset Pricing Model’, *Abacus* 49, 79–81.

Szymanowska, Marta, Frans de Roon, Theo E. Nijman, and Rob van den Goorbergh, 2014, An anatomy of commodity futures risk premia, *Journal of Finance* 69, 453–482.

Telser, Lester G., 1958, Futures trading and the storage of cotton and wheat, *Journal of Political Economy* 66, 233–255.

Titman, Sheridan, K. C. John Wei, and Feixue Xie, 2004, Capital investments and stock returns, *Journal of Financial and Quantitative Analysis* 39, 677.

Tobin, James, 1958, Liquidity preference as behavior towards risk, *Review of Economic Studies* 25, 65.

Treynor, Jack L., 1965, How to rate management of investment funds, *Harvard Business Review* 43, 63–75.

Vassalou, Maria, 2003, News related to future GDP growth as a risk factor in equity returns, *Journal of Financial Economics* 68, 47–73.

Vassalou, Maria, and Yuhang Xing, 2004, Default risk in equity returns, *Journal of Finance* 59, 831–868.

Vayanos, Dimitri, 1998, Transaction costs and asset prices: A dynamic equilibrium model, *Review of Financial Studies* 11, 1–58.

Wallmeier, Martin, and Kathrin Tauscher, 2014, Portfolio overlapping bias in tests of the Fama and French three-factor model, working paper, University of Fribourg.

Welch, Ivo, and Amit Goyal, 2008, A comprehensive look at the empirical performance of equity premium prediction, *Review of Financial Studies* 21, 1455–1508.

White, Halbert, 1980, A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity, *Econometrica* 48, 817–838.

Working, Holbrook, 1948, Theory of the inverse carrying charge in futures markets, *Journal of Farm Economics* 30, 1–28.

Working, Holbrook, 1949, The theory of price of storage, *American Economic Review* 39, 1254–1262.

Zakamulin, Valeriy, 2013, Forecasting the size premium over different time horizons, *Journal of Banking & Finance* 37, 1061–1072.

Zhang, Lu, 2005, The value premium, *Journal of Finance* 60, 67–103.

Zhang, X. F., 2006, Information uncertainty and stock returns, *Journal of Finance* 61, 105–137.

Zhu, Xiaoneng, and Jie Zhu, 2013, Predicting stock returns: A regime-switching combination approach and economic links, *Journal of Banking & Finance* 37, 4120–4133.

Ziegler, Andreas, Michael Schröder, Anja Schulz, and Richard Stehle, 2007, Multifaktormodelle zur Erklärung deutscher Aktienrenditen: Eine empirische Analyse, *Zeitschrift für betriebswirtschaftliche Forschung* 59, 355–389.

Acknowledgments

I extend my sincere appreciation to Prof. Dr. Andreas W. Rathgeber (dissertation advisor and first reviewer), Prof. Dr. Marco Wilkens (second reviewer), Prof. Dr. Hans Ulrich Buhl, Prof. Dr. Siegfried R. Horn, Prof. Dr. Stefan Stöckl, Prof. Dr. Martin Wallmeier, Prof. Dr. John A. Doukas (editor of the journal *European Financial Management*), Prof. Dr. Tarun K. Mukherjee (editor of the journal *Review of Financial Economics*) as well as to two anonymous journal referees for their helpful comments and their support.