Developing an Equity Factor Model for Risk

Daniel Straumann and Tarik Garidi*
RiskMetrics Group
daniel.straumann@riskmetrics.com, tarik.garidi@riskmetrics.com

This paper gives a survey on equity factor modeling and how it relates to other risk modeling techniques. The theory is illustrated by means of an empirical study.

1 Introduction: two common risk modeling approaches

Factor modeling is intimately connected to the notion of risk and return. It is one of the fundamental laws of economics that returns come at the price of risk. While it is rather clear how to measure or define return, there are various definitions and measures of risk. Traditionally, asset management has relied on standard deviations whereas the risk management community prefers Value-at-Risk (VaR) and other tail statistics. Applied to portfolios, these measures summarize the profit and loss (P&L) distribution in terms of risk. The P&L distribution is derived from the joint returns, and therefore accurately modeling them is important. There exists a variety of modeling practices for stock returns, which have dominated different market segments and have influenced the architecture of risk and portfolio management tools in their specific way. We distinguish between two modeling approaches for returns: the current RiskMetrics methodology, which we also call *granular approach*, and the *equity factor model approach*. It is very important to recognize that from a statistical point of view these approaches only differ in the *way* return scenarios are generated. Both approaches have their strengths and weaknesses, and therefore it is our goal to discuss them in the following.

1.1 Granular models

By starting with the granular approach, history is in a sense reversed. Equity factor models became popular in the 1970s whereas the granular approach, very much driven by the publicly available

*The authors would like to thank Prof. J. Bai for interesting discussions and Dr. C. Finger for very careful editing work.

RiskMetrics documents, started to be used in the 1990s.

The granular approach emerged from the need of investment banks to accurately measure, aggregate, and attribute risks, across several trading books and asset classes. Since they traded increasingly complicated instruments and started to pursue long-short strategies, people needed a model for stock returns granular enough to capture, for example, pair trading risk. For this reason, one wanted to analyze securities individually. In this spirit RiskMetrics developed a scalable framework that relies on pricing models and where *each* observed market variable—equity and commodity price, FX and interest rate, credit spread, and implied volatility—*is* a risk driver. We just mention the most important points and refer to *Return to RiskMetrics* (Mina and Xiao 2001) for a more detailed treatment. Each log-return is a risk factor, and log-returns are assumed jointly normal with a stochastic volatility imposed. Derivatives and bonds are simulated through applying the pricing function to scenarios of the underlying risk factors. In this way, non-linearities and to some extent heavy tails are captured. A criticism of the granular approach that we have encountered concerns the simulation of the multivariate normal risk factors. The standard simulation method presented in statistics textbooks is based on a Cholesky decomposition of the sample covariance matrix, which is indeed computationally costly and might be numerically instable for large dimensions. However, by virtue of an elegant mathematical trick one can avoid the computation of the sample covariance matrix and thus simulate risk factors quickly; see (Benson and Zangari 1997).

Summarizing, the granular approach relies on broad market data coverage and availability of pricing models. It is flexible in terms of asset class and market coverage and allows for a granular portfolio risk analysis. Given the increasing portfolio complexity it has become an essential approach for comprehensive risk management.

1.2 Equity factor models

To understand the development of financial factor models, let us first give some historical background. Markowitz (1952) was the first to give a mathematical model for portfolio risk and return. Using volatility as the risk measure, he showed how efficient portfolios can be constructed. One of the lessons from his theory is that portfolio diversification reduces risk. Building on the Markowitz framework and assuming that all investors have the same views, hold efficient portfolios and that there is a risk-free asset,[1] Sharpe (1964) derived what is now known as CAPM (Capital Asset Pricing Model). The CAPM postulates that expected excess returns (above risk-free) of securities (or portfolios) are linearly related to

[1]This means: an asset which has volatility zero.

the excess returns of the market portfolio via the regression coefficient β versus the market and an intercept of zero. Therefore, when returns are decomposed into the market component and an uncorrelated remainder, the CAPM says that the investor is only compensated for bearing the systematic risk, that is being exposed to the market. The remainder part in the decomposition, also called specific risk, does not give any extra return. In other words, the return of a security or portfolio is uniquely determined by the exposure to the market, or rephrased in a more sloppy way, "only the market matters". Since expected excess returns are not directly observed, one introduced the *market model*, a statistically testable version of the CAPM:

$$r_i - r^f = \alpha_i + \beta_i(r^m - r^f) + \varepsilon_i, \tag{1}$$

where r_i and r_m are the realized returns on security i and the market, respectively, r^f is the risk-free rate and the specific returns ε_i have mean zero, are pairwise uncorrelated and uncorrelated with r^m. If the CAPM is valid, clearly $\alpha_i = 0$. The market model is the ancestor of all financial factor models. It contains one factor: the market excess return. Note that the market model is much simpler than the granular model because its correlation structure is determined by the β_i and therefore constrained.

The one-factor model was then generalized by Ross (1976), who proposed to work in a multi-factor setting. In his APT (arbitrage pricing theory) he showed (under certain assumptions) that the factors carry risk premiums, to which the expected return of securities are linearly related via the factor exposures. APT remains however totally silent about how many factors have to be included and how they should be constructed: APT just assumes they exist. At the same time that APT was invented, empirical work and tests of CAPM and APT started to flourish and statistical aspects of financial factor models became investigated. The so-called anomalies literature had a large impact since it found that the market model is not very accurate. As a contradiction to the market model, certain other variables, such as company size, have explanatory power (Banz 1981), and should therefore be included in a factor model. These observations were later synthesized by Fama and French (1993), who came up with what is now known as the Fama-French three-factor model.

With factor modeling, several goals are pursued. A factor model provides a means to reduce the dimension. A further important goal of factor modeling is to identify sources of risk and to decompose risk into systematic and specific (or idiosyncratic) components. This decomposition assists managers in identifying and understanding the nature of risk and in controlling it through diversification and investment decisions. Well constructed factors (such as style, industry, macroeconomic, country) can easily be related to investment strategies, and therefore help portfolio managers making "risk aware" investment decisions, ensure that portfolios are in line with investors views on returns (strategies) and with their risk appetite. Optimization and portfolio allocation techniques within a factor model framework have been developed for these purposes.

Summarizing, the three main goals when constructing a factor model are:

- Be accurate — explain most of the equity variance and co-movements,
- Be parsimonious — use few factors,
- Choose economically meaningful factors in order to allow for an intuitive understanding of the sources of risk to which the portfolio is exposed.

1.3 Comparison of granular model and equity factor model

In factor models, simplicity is achieved by the idea that factors influence each security through a linear relationship, and clarity relies on unambiguous and few factors that allow for a clear interpretation. These features contributed to the popularity of factor models. However, the argument of simplicity and clarity is weakened in the light of the following issues:

- For portfolios that contain derivatives, linearization of the corresponding pricing functions can be applied in order to come up with a factor decomposition. This does not always work very well, particularly over long analysis horizons. Another solution would be to enrich the model with factors which themselves impart on the derivatives nonlinear behavior. In the latter case, modeling the dynamics of these nonlinear factors becomes the crucial problem.
- Factors are usually constructed in order to describe a given asset class, following the idea that similar securities should display similar returns. As a result, cross-asset class portfolios require the combination of many different types of factor models, not necessarily compatible. Moreover the number of factors may grow large, and interpretability may become difficult due to cross correlations.

Note that on the other hand derivatives and complex portfolios are readily handled in the granular approach. It is therefore natural for any risk system to offer *both* modeling approaches.

An important shortcoming of the granular approach is related to optimization and covariance estimation. The factor model contains significantly fewer parameters than the granular model. Indeed, for the granular model describing K stocks to be fully specified, one would have to estimate the sample

Table 1

Large cap portfolio annualized volatility decomposition on 31 January 2005

Results are based on monthly returns and an EWMA parameter of 0.9.

Risk Model	**Total risk**	**Risk decomposition**			
		Market	SMB	HML	Specific
Granular	7.17	xx	xx	xx	xx
Market model	7.24	6.84	xx	xx	2.15
Fama-French model	7.30	7.91	1.64	0.88	1.59

covariance matrix, which contains $K(K+1)/2$ unique entries and the K means for the securities. In contrast, in a factor model with p factors, one needs to estimate $(p+1)$ coefficients per security together with one specific variance. Moreover, the factor covariances and means have to be estimated. This gives $K(p+2)+p(p+3)/2$ parameters, which is small compared to the number of covariances. For example, a portfolio of 100 stocks would require 5150 parameters to be estimated in the granular approach as opposed to only 720 for a five-factor model.

The large number of parameters in the granular model can lead to problems in covariance matrix estimates. Note that if the number of return observations is strictly smaller than the number of assets, then the sample covariance matrix is singular. This creates troubles for optimization. By contrast, the covariance matrix estimate implied by the factor model is always non-singular. Moreover, as has been shown in (Fan, Fan, and Lv 2006), this estimator has better asymptotic properties than the sample covariance matrix.[2]

Our goal in adding a factor model is not to produce better risk figures, but to provide insight on risk characteristics of portfolios. Indeed, the total risk for a given security or portfolio should ideally not depend on the model that is used, and one does expect similar results.

In Table 1, we present the annualized volatility for an equity portfolio consisting of 53 US large cap stocks, calculated using two different factor models and the granular approach. The first thing to note is that although different models lead to different results, the total portfolio risk is similar for all three risk models. Factor modeling does not add a lot to the granular approach if only the total risk is needed. The

[2]This statement is only true provided the factor model has been *correctly* specified, this means, the underlying risk factors are known and observed.

Table 2
Large cap portfolio factor exposures on 31 January 2005
Results are based on monthly returns and an EWMA parameter of 0.9.

Risk Model	**Exposures**			
	Alpha	Market	SMB	HML
Granular	xx	xx	xx	xx
Market model	-0.01	0.80	xx	xx
Fama-French model	0.05	0.93	-0.19	-0.18

true valued added by the factor model lies in its capability to decompose the portfolio risk. In our example, the market model describes the risk in terms of a market factor and a specific part, and the Fama-French model adds the size (SMB) and value/growth (HML) factors. Exposures of the portfolio with respect to these risk dimensions are shown in Table 2. In contrast to granular models, factor models provide a natural framework for calculating exposures. Knowing exposures, one can manage a portfolio according to particular views on factors.

Our goal for this article is to give an overview of the various frameworks that can be used to implement an equity factor model, and to present our model and some of our factors through an empirical study based on a test portfolio framework. In Section 2, we give definitions and notations for linear factor models and discuss the factors themselves, the various types and construction techniques. Section 3 is dedicated to an empirical study. Factors are useful only if they provide intuition and if they have explanatory power. We will illustrate the theory and assess the constructed factors.

2 Linear factor models

This section provides a comprehensive treatment of equity factor models. It introduces mathematical notation and definitions, and discusses the taxonomy of factors and construction methods.

2.1 Definition

There are many possible definitions of a linear factor model. To begin with, let us introduce some notation. We are given a universe of stocks, indexed by $i = 1, \dots, K$, and our goal is to model their joint price evolution in discrete time. To this end, we look at *relative* stock returns

$$r_i(t) = \frac{p_i(t) - p_i(t-1)}{p_i(t-1)} = \frac{p_i(t)}{p_i(t-1)} - 1. \tag{2}$$

Here $p_i(t)$ stands for the ith share price at time t, or more precisely, the share price adjusted for dividends and splits. We assume that dividends are instantly reinvested in the same stock, and therefore $r_i(t)$ is a *total* return over the *return period* $(t-1, t]$. Often people work with excess returns, that is, returns where the risk-free rate is deducted. In what follows, returns are assumed to be total excess returns unless explicitly stated otherwise.

Note that as opposed to the RiskMetrics methodology,[3] we do not work with the log-returns $\log(p_i(t)/p_i(t-1)) = \log(1 + r_i(t))$. The reason for this choice is that relative returns are linear with respect to aggregation over stocks, whereas log-returns are nonlinear. For the asset management world, which primarily deals with portfolio analysis, linearity of returns when forming portfolios is a necessity, and hence relative returns are usually preferred. The cost for using relative returns in that they are bounded from below by -1, and therefore should realistically be modeled by asymmetric distributions,[4] whereas log-returns can be regarded as symmetric. Note that it is always possible to go back and forth from relative returns to log-returns via a one-to-one mapping, thereby exploiting the advantages of the respective return formats. Moreover, a Taylor argument shows that the log-return coincides with $r_i(t)$ up to first order.

Now our goal is to model the stock relative returns $r_i(t)$. For our needs of describing joint stock returns,

[3] Presented in *Return to RiskMetrics* (Mina and Xiao 2001)

[4] To the best of our knowledge, starting from (Markowitz 1952) the literature mostly makes the (tacit) assumption that relative returns are Gaussian. This can be justified provided the parameters are such that the probability of falling below -1 is negligible.

a linear factor model consists of three elements:

1. A finite number of factor time series $\{f_j(t) \mid j = 1, \ldots, p\}$,

2. A linear link between returns and factors, namely

$$r_i(t) = \alpha_i(t) + \sum_{j=1}^{p} \beta_{ij}(t) f_j(t) + \varepsilon_i(t), \qquad i = 1, \ldots, K, \tag{3}$$

   where the *idiosyncratic* (or specific) returns $(\varepsilon_1(t), \ldots, \varepsilon_K(t))$ are uncorrelated with mean zero, and uncorrelated with $\{f_j(t)\}$, simultaneously for all t. The regression coefficients $\beta_{ij}(t)$ are also referred to as exposures or factor loadings. Note that α_i and β_{ij} are time-dependent, so strictly speaking, we consider *conditional* factor models; see (Ang and Chen 2005).

3. A model for the dynamics of $\varepsilon_i(t)$ and $f_j(t)$.

A few words about the latter mathematical definition are necessary. First, observe that the definition remains silent about which and how many factors to include. Equation (3) is merely a hypothesis about the truth of stock returns comovements, and the challenging task is left to the econometrician to propose factors and verify in data samples that the relationship (3) is a reasonable assumption.

Economically the f_j represent common sources of risk. The term $\alpha_i(t) + \sum_{j=1}^{p} \beta_{ij} f_j(t)$ is also referred to as systematic return, because in contrast to the specific return it cannot be diversified away by forming portfolios of stocks.[5] Ross (1976) showed in his celebrated APT theorem, later adapted to the case of finitely many assets in (Chen and Ingersoll 1983), that in linear factor models there exist $\lambda_j(t)$ such that $\mathrm{E}(r_i(t)) = r^f(t) + \sum_{j=1}^{p} \beta_{ij}(t)\lambda_j(t)$, where $r^f(t)$ is the risk-free rate; the $\lambda_j(t)$ can be regarded as risk premiums. As a consequence of this result, a lot of empirical research concentrated on finding so-called priced factors, that is, factors f_j with $\mathrm{E}(\lambda_j(t)) > 0$. Since we are interested in explaining risk well more than in pricing the factors, we do not require that our factors carry non-zero risk premia.

Finally, for building benchmarks it would be desirable that the factors are *investable*, which means that they are the returns on portfolios of investable assets, However, since we seek to find models with high explanatory power, we take the pragmatic approach of also admitting non-investable factors.

Note that the three items above uniquely determine the stochastic process of joint stock returns. If one is only interested in a descriptive and backward-looking framework for stock returns, the third element is

[5] Diversifiability is implied by the uncorrelatedness of the $\varepsilon_i(t)$.

not needed. However, for meaningful risk or return forecasts, the time evolution of factors and idiosyncratic returns must be modeled. The majority of the literature on factor models deals with the static case of finding good factors and remains silent about the factor dynamics. Since the first step in constructing a factor model always consists of obtaining reasonable factors, we align ourselves with the literature and concentrate on discussing and evaluating factors. We deal with the static case and assume that constructed factors follow the RiskMetrics dynamic model. Future work will address factor dynamics.

From equation (3) and the assumption that idiosyncratic returns are uncorrelated with the factors, we can immediately express return covariances by their loadings and the factor covariances:

$$\mathrm{Cov}\left(r_i(t), r_{i'}(t)\right) = \sum_{j=1}^{p} \sum_{j'=1}^{p} \beta_{ij}(t)\beta_{i'j'}(t)\,\mathrm{Cov}\left(f_j(t), f_{j'}(t)\right) + \mathrm{Cov}\left(\varepsilon_i(t), \varepsilon_{i'}(t)\right). \tag{4}$$

Our derivation assumes that the $\alpha_i(t)$ and $\beta_{ij}(t)$ are deterministic. Moreover, note that this formula was implicitly used earlier in this paper when we derived the number of free parameters in a factor model. An important quantity in the context of factor model is the so-called measure of determination, commonly known as R^2. The R^2 is the proportion of variance explained by the systematic component with respect to the total variance. More formally:

$$R_i^2(t) = \frac{\mathrm{Var}\left(\sum_{i=1}^{p} \beta_{ij}(t) f_j(t)\right)}{\mathrm{Var}(r_i(t))} = 1 - \frac{\mathrm{Var}\left(\varepsilon_i(t)\right)}{\mathrm{Var}(r_i(t))}. \tag{5}$$

The higher the R^2, the better are the returns explained by the factors. If R^2 was one, the factors would perfectly predict the returns. In the following two sections, we present various techniques that can be used to specify the models: construct, classify the factors, estimate the residuals (this means the idiosyncratic returns), and determine the loadings. First, we will focus on the factors.

2.2 Factor taxonomy

In this section we would like to give an overview of the different factor types and factor construction methods. In a first step, in order to fix the terminology we give a classification, which is similar to the ones given in (Connor 1995; Chan, Karcoski, and Lakonishok 1998; Zangari 2003).

Equity factors represent certain economic sources of risk. Our classification separates according to the ingredients used to construct the factors. We distinguish between six types of factors:

Macroeconomic. There is a wide consensus that stock market performance is affected by the state of the economy, and this has also been shown empirically; see (Chen, Roll, and Ross 1986). Macroeconomic factors are derived from macroeconomic variables, interest rates, FX rates, and commodity prices. Common examples of macroeconomic factors are:

- relative oil price changes,
- changes in the average default premium, as measured by the difference between returns of a high-yield bond index and the return on long-term government bonds,
- changes in the maturity premium, as measured by the difference between the return on long-term government bonds and the one-month Treasury bill return,
- changes of growth rate of the GDP (gross domestic product),
- changes of the consumer CPI (consumer price index).

Macroeconomic variables are problematic insofar as they are mostly reported with a certain time lag, so that their use in real time has its limitations.

Market. Originally motivated by the CAPM, these factors stand for the general market risk. For US stocks, one could for instance use S&P 500, Russell 3000 or Wilshire 5000 index returns as the market factor. Having a historical price database at hand, one can easily create market factors by computing the returns of the value-weighted portfolio consisting of all US stocks.[6]

Sector/Industry. These factors represent sources of risk related to economic sectors or industries. Note that sectors are groups of industries, and therefore there are more industry factors than sector factors. Industry factors will allow a more granular analysis. Examples are the MSCI sector/industry returns. The MSCI sector/industry indices rely on the GICS (Global Industry Classification Standard), a standard which was jointly defined by Standard & Poors and Morgan Stanley.

Fundamental. Fundamental factors are factors derived from historical stock-specific fundamentals, together with the corresponding returns. Fundamentals which are often used are the market capitalization, which results in so-called size factors, and the price-to-book ratio, which yields value/growth factors. There are a multitude of fundamental factors which apparently have certain explanatory power; see (Stephan, Maurer, and Dürr 2000).

[6] Here one has to specify what is meant by “all US stocks”, this means one has to define the so-called “universe”; see Section 3.1.1 below.

Technical. Technical factors are constructed using historical, stock specific, non-fundamentals data. Examples include liquidity measures or a firm's past return data which have been found to explain future returns; see (Chan, Jegadeesh, and Lakonishok 1996).

Statistical. In the equity factor model terminology, statistical factors are the factors extracted from historical returns, via principal components analysis (PCA) or factor analytical methods. The notion "statistical factors" is perhaps a bit unfortunate because it might lead to the idea that they are the only factors constructed by means of statistical methods. This is however not the case: many of the factors previously described can also be constructed by means of cross-sectional regressions, as will be discussed in more detail below.

In the following, we will distinguish between *predefined* and *estimated* factors. We say that a factor is *predefined* if it comes from the returns of a trading strategy or if it is observed, that is, if its value is available or easily derivable from market or other data. To illustrate this definition, consider the following examples of predefined factors. MSCI returns would be observable factors (and thus also predefined), whereas a market factor calculated from the value-weighted portfolio of US stocks would be predefined, but not observable. In the case of *estimated* factors, statistical methods such as PCA or cross-sectional regressions are used to construct them.

It would be wrong to take this classification literally and to believe that the different types do not overlap. It might well be that a certain factor is highly correlated to the linear combination of factors belonging to a different class. It is merely a matter of the portfolio manager's taste and goals which factor types he wants to use. As an example, a manager who is particularly interested in understanding the portfolio risk in terms of exposures to industries will take industry factors rather than macroeconomic factors. Under this aspect we believe that for a successful equity factor model flexibility and customizability is key.

We have introduced a terminology for distinguishing the various factors. Next we show how loadings are estimated from predefined factors. This will be followed by a taxonomy of factor construction methods.

2.3 Time series regression

Here we describe how one can compute the factor loadings from predefined factors. We appeal to equation (3) for the stock returns,

$$r_i(t) = \alpha_i(t) + \sum_{j=1}^{p} \beta_{ij}(t) f_j(t) + \varepsilon_i(t), \qquad i = 1, \ldots, K,$$

and we assume that the past and present factor values f_j are given. Then a natural way to obtain the loadings $\beta_{ij}(t)$ at a certain time t is *time series regression*. This means, for each stock i, take the returns on the time window $(t-m,t]$ and regress them on the factors f_j. Since our empirical study in Section 3 heavily relies on time series regressions, it is appropriate to be a bit more explicit and to provide formulas. To this end, introduce the $m \times 1$-vectors

$$\mathbf{r}_i(t) = \begin{pmatrix} r_i(t-m+1) \\ \vdots \\ r_i(t) \end{pmatrix}, \mathbf{f}_j(t) = \begin{pmatrix} f_j(t-m+1) \\ \vdots \\ f_j(t) \end{pmatrix}, \varepsilon_i(t) = \begin{pmatrix} \varepsilon_i(t-m+1) \\ \vdots \\ \varepsilon_i(t) \end{pmatrix} \tag{6}$$

and the $m \times (p+1)$ design matrix of factors

$$\mathbf{F}(t) = (\mathbf{1}_{m\times 1}, \mathbf{f}_1(t), \ldots, \mathbf{f}_p(t)), \tag{7}$$

where $\mathbf{1}_{m\times 1}$ denotes the $m \times 1$-matrix of ones. Note that our notation suppresses the window size m. For estimation purposes we suppose that $\alpha_i(t) = \alpha_i$ and $\beta_{ij}(t) = \beta_{ij}$. Then, as a consequence of our assumptions for the linear factor model, we have that

$$\mathbf{r}_i(t) = \alpha_i + \mathbf{F}(t)\beta_{i\cdot} + \varepsilon_i(t), \tag{8}$$

with $\beta_{i\cdot} = (\beta_{i1}, \ldots, \beta_{ip})'$. The ordinary least-squares (OLS) estimate[7] applied to the latter equation gives the estimators $\widetilde{\alpha}_i(t)$ and $\widetilde{\beta}_{i\cdot}(t)$, that is,

$$\begin{pmatrix} \widetilde{\alpha}_i(t) \\ \widetilde{\beta}_{i\cdot}(t) \end{pmatrix} = (\mathbf{F}(t)'\mathbf{F}(t))^{-1}\mathbf{F}(t)'\mathbf{r}_i(t). \tag{9}$$

Refinements of this estimator are possible. First, OLS is recommended only if the covariance matrix of the past and present idiosyncratic returns is a multiple of the unit matrix, that is, $\mathrm{Cov}(\varepsilon_i(t)) = \sigma_i^2 I_{m\times m}$. It is however often the case that the idiosyncratic returns are heteroscedastic, that is, they have a time-varying variance. To increase the efficiency of the estimator for the loadings, one is better off using a two step procedure, where in a first step the variances of the residuals are estimated, and in the second step the final parameters. This could look as follows:[8]

[7] Recall that OLS minimizes

$$|\mathbf{r}_i(t) - \alpha_i + \mathbf{F}(t)\beta_{i\cdot}|^2 = \sum_{\ell=1}^{m} \left(r_i(t-m+l) - \alpha_i - \sum_{j=1}^{p} \beta_{ij} f_j(t-m+\ell) \right)^2.$$

[8] See (Northfield 2006).

1. Estimate the idiosyncratic returns $\varepsilon_i(t)$ by means of OLS:

$$\widetilde{\varepsilon}_i(t) = \mathbf{r}_i(t) - \widetilde{\alpha}_i(t) - \mathbf{F}(t)\,\widetilde{\beta}_{i\cdot}(t). \tag{10}$$

   Then $\overline{\sigma}_i^2 = (m-p)^{-1}\sum_{k=1}^{m}\widetilde{\varepsilon}_i(t-m+k)^2$ can be seen as the mean level of the residual variance, and we set

$$\widetilde{\sigma}_i^2(t-m+k) = \max(\overline{\sigma}_i^2, \widetilde{\varepsilon}_i(t-m+k)^2). \tag{11}$$

2. Apply generalized least-squares estimation (GLS) to the linear model (8), using the diagonal covariance matrix

$$\widetilde{\Omega}_i(t) = \operatorname{diag}\left(\widetilde{\sigma}_i^2(t-m+k), \dots, \widetilde{\sigma}_i^2(t)\right). \tag{12}$$

   for $\varepsilon_i(t)$. This gives the estimator

$$\begin{pmatrix} \widehat{\alpha}_i(t) \\ \widehat{\beta}_{i\cdot}(t) \end{pmatrix} = (\mathbf{F}(t)'\widetilde{\Omega}_i(t)^{-1}\mathbf{F}(t))^{-1}\mathbf{F}(t)'\widetilde{\Omega}_i(t)^{-1}\,\mathbf{r}_i(t). \tag{13}$$

   As a matter of fact, GLS is just OLS applied to the transformed data $\widetilde{\Omega}_i^{-1/2}\mathbf{r}_i(t)$ with a transformed factor matrix $\widetilde{\Omega}_i^{-1/2}\mathbf{F}_i(t)$. From this one can also interpret GLS as *weighted* least-squares with weights equal to $1/\widetilde{\sigma}_i^2(t-m+k)$. Note that due to the special form of the variance estimated in the first step, the observations with a residual *above* the level $\overline{\sigma}_i$ are given less weight.[9]

The second improvement is concerned with the time dependency of the coefficients $\alpha_i(t)$ and $\beta_{ij}(t)$. For estimation purposes we assumed fixed levels, but it is important to note that the estimators $\widehat{\alpha}_i(t)$ and $\widehat{\beta}_{ij}(t)$ are time-varying because the time window over which the regressions are performed moves forward in time; therefore some of the time-variation is captured. A proper estimation is only possible when more structure on the $\alpha_i(t)$'s and $\beta_{ij}(t)$'s is added. The commonly taken approach is based on Kalman filtering techniques; see (Swinkels and van der Sluis 2006). As a simple ad-hoc approach, which yields an estimator that reacts quicker to changes of the market conditions, we propose to weight the residuals exponentially. This is achieved by applying the OLS or two-step estimators to returns $\mathbf{r}_i(t)$ and factors $\mathbf{F}(t)$ initially transformed by the $m \times m$ weight matrix

$$W = \operatorname{diag}(\lambda^{(m-1)/2}, \lambda^{(m-2)/2}, \dots, 1). \tag{14}$$

[9] This means, one minimize the objective function

$$\sum_{\ell=1}^{m}\frac{1}{\widetilde{\sigma}_i^2(t-m+k)}\left(r_i(t-m+l) - \alpha_i - \sum_{j=1}^{p}\beta_{ij}f_j(t-m+\ell)\right)^2.$$

Here $0 < \lambda \leq 1$ is an appropriately chosen decay rate. Note also that applying exponential weights leads to a nice property:if sample covariances are calculated by means of EWMA[10] with identical λ, then the sample covariance between the factors and the residuals is zero, as desired.

2.4 Factor construction

One has to distinguish between two cases: the factor is *observed* or the factor is *unobserved*. In the first case, there is of course nothing to do. To the best of our knowledge, macroeconomic factors are mostly observed, albeit macroeconomic variables are often reported with a time lag and not at the necessary observation frequency (for example GDP growth). If the factors are to be used in real-time, as a way to overcome this problem. one might fit a dynamic model to past data (for instance an autoregressive integrated process) and then try to use forecasted values as a proxy for the present, but not yet available factor realizations. Recall that market and industry/sector factors based on indices (such as MSCI) are also classified as observed.

We now turn to unobserved factors, that is, factors that result from calculations. Recall that the goal is to express historical factor values from past and present stock returns and possibly firm attributes such as fundamentals. We broadly distinguish between three construction methods: factor-mimicking portfolios (FMP), cross-sectional regressions, and PCA. The cross-sectional regressions and PCA approach are statistical methods to derive *estimated* factors. On the other hand, FMP factors are predefined and require no other computations than return manipulation (sums, differences, portfolio returns). Factor types and construction factor types are summarized in Figure 1.

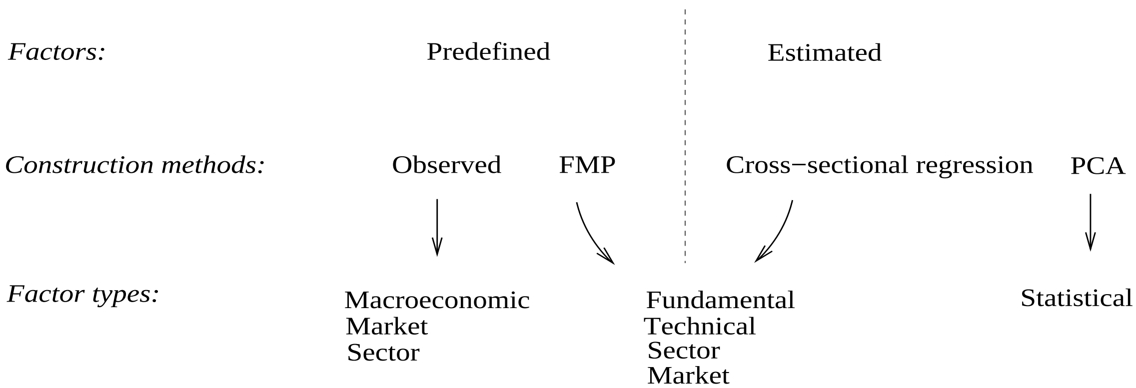
2.4.1 Factor-mimicking portfolios

The term factor-mimicking portfolio comes from the idea that factors can be represented or proxied by returns of appropriately defined portfolios. The simplest example is the market factor. Indeed, the market factor is mostly defined as the return of the *market portfolio*, which is just the value-weighted return of all stocks in the universe. If $c_i(t)$ denotes the market capitalization of stock i at time t, then the market factor is calculated as

$$f^{Mkt}(t) = \sum_{i=1}^{K} w_i(t-1) r_i(t) \ \text{ with } \ w_i(t) = \left(\sum_{k=1}^{K} c_k(t) \right)^{-1} c_i(t). \tag{15}$$

[10] See (Mina and Xiao 2001).

Figure 1
Factor construction methods and related factors



Note that using weights proportional to market capitalizations observed at the beginning of the return period $(t-1,t]$ makes the market factor investable. Another example are sector or industry factors, which are just the returns of the value-weighted portfolio of all the stocks belonging to a certain sector or industry.

So far we have discussed long-only factor-mimicking portfolios. Another type of factors has been introduced by Fama and French (1993). Due to its simplicity and elegance, their idea has had a significant impact on portfolio management and financial economics. They consider the returns of long-short portfolios, formed according to a sort with respect to a certain stock attribute. Let us explain their SMB (small minus big) factor, sometimes also called the size factor, which appears to carry the most explanatory power. At each portfolio *reconstruction date* (say end of June every year), all stocks in the universe are ranked according to market capitalization. Let A consist of the stocks that fall in the lower 50% by market cap and B of the stocks in the 50 % upper. Apply value weighting within each portfolio. Then the SMB factor is just the difference of portfolio returns $r^A(t)$ and $r^B(t)$, that is,

$$f^{SMB}(t) = r^A(t) - r^B(t). \tag{16}$$

The SMB proxies for a factor related to firm size. It is basically a return spread between small and large companies and therefore shows how much the small companies outperform (or underperform) large companies in a certain return period. Some regard $f^{SMB}(t)$ as a *style factor*. Indeed, an investor who pursues a small cap style is heavily exposed to f^{SMB} (and the market factor f^{Mkt}). However, this statement is only accurate for large portfolios. For individual stocks it is well possible that a large cap has a *positive* exposure to SMB or a small cap has a *negative* exposure to SMB. This would then mean that the large cap stock does well in an economic environment that is favorable for the small cap segment, and vice versa.

The Fama-French principle for constructing factors can be applied to any kind of attribute (or combinations of attributes). If the stocks are ranked according to the price-to-book ratio,[11] the so-called Fama-French HML (high minus low) or value/growth factor is obtained. This is the performance of value stocks (low price-to-book ratio) relative to growth stocks (high price-two-book ratio).[12] Both SMB and HML are fundamental factors. It is also possible to construct technical factors. As an example, the momentum factor comes from using as an attribute the return over a past period. The liquidity factor is constructed by using some liquidity measure over a past period.

To illustrate the use of the Fama-French factors, we graph in Figure 2 the factor exposures of General Electric (GE) in the three-factor model, which includes the market factor together with SMB and HML. The results are not surprising: GE has a market β which oscillates around one, and GE is strongly negatively related to SMB, which is in line with the fact that the company is among the three largest companies in the world. Moreover, the negative loading with respect to HML shows that the market perceives GE as a growth company. In view of the fact that GE's mean price-to-book ratio over the observation period was 6.5 (compared to 3.0 for the average US company), this is not astonishing either.

We have suppressed certain practical issues for computing the factors, such as how the universe is determined, missing data, reconstruction, rebalancing, and so on. This will be covered in greater detail in Section 3. It is important to mention that we have slightly simplified the construction of SMB and HML. In fact, (Fama and French 1993) sort according to *two* dimensions. For SMB they define six value-weighted portfolios formed on size and price-to-book, and take the average return on the three small portfolios minus the average return on the three large portfolios,

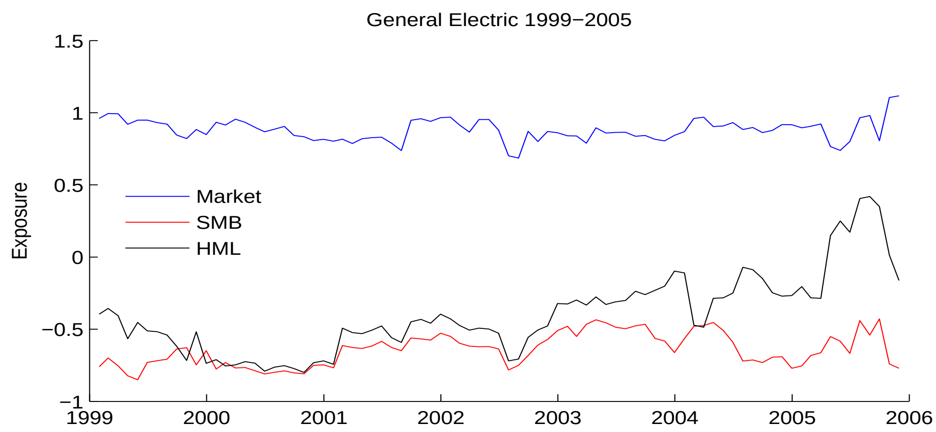
$$\begin{aligned} f^{SMB} &= \frac{1}{3}\left(\text{Small Value} + \text{Small Neutral} + \text{Small Growth}\right) \\ &\quad -\frac{1}{3}\left(\text{Large Value} + \text{Large Neutral} + \text{Large Growth}\right). \end{aligned} \tag{17}$$

The idea behind this correction is to give the value/neutral/growth segments within the small caps or large caps, respectively, equal weights.

[11] The book value is an accounting based attribute found in the balance sheet and the price is the stock price observed on the market.

[12] (Fama and French 1993) actually use book-to-price, and that is why the value/growth factor carries the name HML. Our definition, motivated by the fact that rather price-to-book is reported on financial markets, would justify the name LMH. Another detail worth to be mentioned is the fact that Fama and French look at the ratio lagged for six months.

Figure 2
Factor exposures for General Electric
Exposures are calculated with monthly returns in windows of length two years. The decay factor for the time series regression is set to $\lambda = 1$.



Likewise, the factor HML is defined as the average return on the value portfolios of small and large stocks minus the average return on the corresponding growth portfolios:

$$f^{HML} = \frac{1}{2}(\text{Small Value} + \text{Large Value}) - \frac{1}{2}(\text{Small Growth} + \text{Large Growth})\,. \tag{18}$$

Here one wants to avoid that largely capitalized stocks dominate the HML factor. In Section 3.1.2 we argue that these differences in the definition of SMB and HML do not matter from a statistical point of view.

2.4.2 Cross-sectional regressions

Cross-sectional regression belongs to the *estimated* factor construction type. In contrast to time series regression, one supposes that the loadings $\beta_{ij}(t)$ are known. In comparison to the time-series regression approach from before, the roles of the loadings and factors are interchanged. For every fixed time t, the factor realizations $f_j(t)$ are estimated in terms of the explanatory variables $\beta_{ij}(t)$. Note that this is achieved by means of *one* multiple regression. Indeed, introducing the cross-sectional vectors

$$\mathbf{r}^{CS}(t) = \begin{pmatrix} r_1(t) \\ \vdots \\ r_K(t) \end{pmatrix}, \quad \mathbf{f}^{CS}(t) = \begin{pmatrix} \alpha(t) \\ f_1(t) \\ \vdots \\ f_p(t) \end{pmatrix}, \quad \varepsilon^{CS}(t) = \begin{pmatrix} \varepsilon_1(t) \\ \vdots \\ \varepsilon_K(t) \end{pmatrix} \tag{19}$$

and the cross-sectional $K \times (p+1)$ design matrix

$$\mathbf{B}^{CS}(t) = \begin{pmatrix} 1 & \beta_{11}(t) & \cdots & \beta_{1p}(t) \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \beta_{K1}(t) & \cdots & \beta_{Kp}(t) \end{pmatrix}, \tag{20}$$

we can summarize (3) by

$$\mathbf{r}^{CS}(t) = \mathbf{B}^{CS}(t)\mathbf{f}^{CS}(t) + \varepsilon^{CS}(t). \tag{21}$$

Note that we imposed $\alpha_i(t) = \alpha(t)$ for all i because otherwise the linear system would be over parameterized. The OLS estimator of $\mathbf{f}^{CS}(t)$ is then given by

$$\widehat{\mathbf{f}}^{CS}(t) = \left(\beta(t)'\beta(t)\right)^{-1}\beta(t)'\,\mathbf{r}^{CS}(t). \tag{22}$$

Note that (22) is a weighted sum of returns, and therefore the cross-sectional factors are up to scale of FMP form; see Zangari (Zangari 2003) for further elaboration. Similarly to before, the estimation methodology can be refined, but we omit details.

The loadings $\beta_{ij}(t)$ are certain company attributes such as fundamentals, technical indicators or memberships in a well-defined group of stocks (for instance industry membership). Often one takes the values observed at the beginning of the return period $(t-1,t]$ to calculate $\beta_{ij}(t)$. Examples for attributes are the logarithmic market capitalization, leading to the cross-sectional version of the size factor or the price-to-book ratio, resulting in a value/growth factor. In practice each loading is standardized to have mean zero and unit variance (across the stocks) so that exposures across the different attributes become better comparable. Another method is to combine several attributes into so-called risk indices; see (Stephan, Maurer, and Dürr 2000) for a nice case study applying this method.

As a matter of fact, literature often calls models estimated through cross-sectional regression "fundamental models", but we avoid following this terminology: for us it is rather the data used in the factor construction which determines the factor type, and not the estimation methodology applied. We consider the Fama-French size or value/growth factors already described also as fundamental factors.

Constructing a good factor model by means of the cross-sectional regression method is challenging: one must choose the “right” company attribute, and possibly transform it in order to increase its explanatory power. Working with predefined factors and time series regressions by no means make the tasks easier. Methodologically, it consists merely of changing the focus from loadings to factors.

The differences between the cross-sectional construction method and the factor-mimicking portfolio approach described above are of a more practical nature. With the cross-sectional approach, the loadings need not be estimated. Once $\widehat{\mathbf{F}}^{CS}(t)$ has been evaluated, one can quickly determine the factor decomposition of any equity, including newly listed companies, provided its relevant attributes are known. Since the loadings $\beta_{ij}(t)$ are given at time t, a model with cross-sectional regression factors can take up changes of companies quickly.

From our point of view, the main deficiency of the cross-sectional regression approach lies in the fact that each time a new factor (that is, a new exposure) is added, the past values of the already existing factors change. This is because the factors are estimated *jointly* by means of multiple linear regression, and hence if *one* explanatory variable is added, the coefficients of all the initial explanatory variables not orthogonal to that added variable change. We consider this a very unattractive property. It makes historical comparisons of factors in a model that is from time to time augmented difficult to interpret. Moreover, from a methodological point of view observed factors (such as macroeconomic factors) do not fit into the cross-sectional regression framework. Therefore mixing observed factors and factors obtained from cross-sectional regression might be questionable. Last but not least, we feel that the time series regression with predefined factors is more intuitive and easier to explain (because factors are not estimated). All these considerations prompt RiskMetrics to choose predefined factors, either observed or implied by factor-mimicking portfolios.

2.4.3 Principal component analysis

The third construction technique is PCA. As already mentioned, statistical factors are intimately connected to principal component analysis. PCA is a classical statistical dimension reduction technique. It is well-suited to determine the achievable explanatory power with a certain number of factors since PCA factors are designed to explain variance optimally. Moreover, for a model consisting of predefined factors, PCA might be used to model the residuals, in case some pervasive factor has been missed.

In the context of correlated returns, the idea is to find a small number of uncorrelated linear return combinations that account for most of the variability in the return data. We do not want to go into great detail in this paper and rather refer to Chapter 9 of (Tsay 2005). We just mention that PCA relies on the diagonalization of the $K \times K$ sample covariance matrix $\widehat{\Sigma}$ of the vectors $\mathbf{r}^{CS}(t)$. Let $\widehat{\gamma}_j$ be the eigenvectors of $\widehat{\Sigma}$ with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_K$. Then the jth factor is just the jth principal component,

$$\widehat{f}_j(t) = \widehat{\gamma}'_j \mathbf{r}^{CS}(t). \tag{23}$$

Sometimes the returns $\mathbf{r}^{CS}(t)$ are centered or even standardized in order that very volatile stocks do not dominate the PCA. Since $\widehat{f}_j$ is a linear combination of returns, it can be seen as the return of a portfolio with constant weights given by the coordinates of $\widehat{\gamma}_j$. Since scaling $\widehat{\gamma}_j$ does not essentially change the factor, one can always achieve that these weights sum to one or zero. There are K statistical factors, but the idea is to employ only the first few in a factor model. One can show that the m first factors explain$(\sum_{j=1}^{m} \lambda_j)/(\sum_{j=1}^{K} \lambda_j)$ of the variance in the data.

A variant of PCA is so-called asymptotic PCA, which goes back to (Connor and Korajzcyk 1986). If there are thousands of assets, it becomes quite costly to compute the eigenvalues and eigenvectors of the sample covariance matrix. To circumvent the problem, one can invert the role of the time index t and the asset index i. The resulting sample covariance matrix has a dimension equal to the number of observed return periods, and its eigenvectors $\widetilde{f}_j(t)$ serve as factors. It can be shown that each factor can be represented as $\widetilde{f}_j(t) = \widetilde{\gamma}'_j \mathbf{r}^{CS}(t)$, and each $\widetilde{f}_j(t)$ is up to a multiplicative constant equal to $\widehat{f}_j(t)$. This shows that asymptotic PCA should be regarded as a numerical trick. We made use of it for our empirical study in Section 3.

The main problem with statistical factors is to give them an intuitive interpretation. It is a stylized fact that the first principal component represents the market, but this is generally the only factor that is easily interpretable. Moreover, when re-estimating the PCA factors with a new data sample, it can happen that factors with eigenvalues close to each other are permuted, which is undesirable.

3 Empirical study

The remainder of this article describes the RiskMetrics framework for constructing and assessing equity factors. We are addressing questions like:

- How relevant are factors ?
- How many factors should be employed to describe a portfolio ?
- How should factors be selected to obtain a good model ?

As we have alluded to already in Section 2, our framework is based on predefined factors because we want to achieve a high degree of flexibility and generality. We employ market, sector, fundamental and technical factors and construct them by factor-mimicking portfolios.

Model assessment is tightly knit with the goals of factor modeling: explaining returns well through few factors, which are moreover economically interpretable. A natural measure for the *explanatory* power of factors is the *sample* R^2, which is the statistically estimated counterpart of (5). In outputs of regressions, the sample (or regression) R^2 is often the first number people look at, and the goal is to achieve high values. Following the common statistical language, we will not distinguish between sample R^2 and R^2 in the following.

There is a tradeoff between high explanatory power and parsimony. As a matter of fact, by the definition of the R^2, adding factors to a model *automatically* increases its value, and therefore it is always possible to achieve values close to one when sufficiently many explanatory variables are included. This conflicts, however, with the requirement of parsimony. Correlations between factors may furthermore dilute interpretability. From a statistical point of view, a high R^2 does not necessarily mean that the fitted linear model is accurate. It could simply mean that the model overfits the data, which results in low *predictive* power of the model. The R^2 is by no means a measure for the goodness of a model, and its interpretation can be delicate. These aspect will be further discussed on the basis of practical examples.

Another aspect not yet mentioned is the question against which securities or portfolios the factors should be tested. This is related to the intended use of the factor model. The choice of a testing environment determines to a large extent the specification and selection of factors. It is rather clear that a manager of a large cap fund has totally different needs with respect to factors from a gold fund manager. RiskMetrics has developed a testing framework based on randomly chosen general portfolios, which we will discuss further below.

We have the feeling that the above points, although fundamental and simple, are often not explicitly expressed in the mainstream literature. Many papers deal with the problem of presenting and defending a new factor candidate, and by testing it against specifically adapted portfolios, the factors's explanatory power is established. The question of predictive power is seldom addressed.

In the following, we describe in more detail the factors that we construct and present our testing framework and the results.

3.1 Factor description

The factors that we use in our study are seventeen equity factors: the market factor, six style factors (size (SMB), value/growth (HML), momentum, dividend yield, liquidity, volatility), and ten industry/sector factors based on the standard GICS classification. All factors are derived from corresponding factor-mimicking portfolios as introduced in Section 2.4.1. In order to be as transparent as possible, and to facilitate interpretability of the factors, it is necessary to provide details on the input data and mechanics used for forming the FMPs.

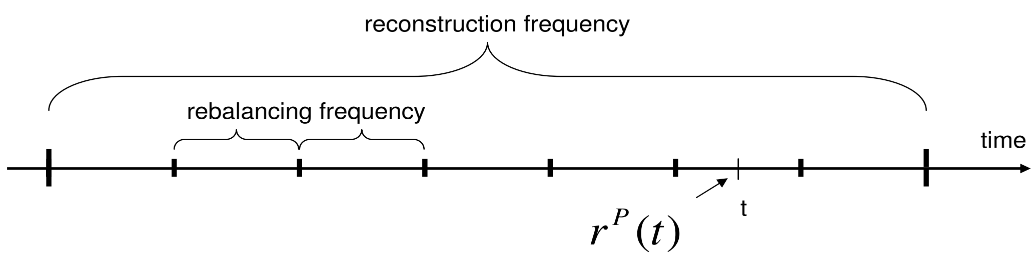
3.1.1 Construction method of factor mimicking portfolios

We have used eleven years (January 1995 to December 2005) of monthly Compustat data for both fundamental and price/return information. A filtering algorithm is applied to all US active and inactive companies in order to construct a representative subset called the universe. We base the universe on primary issues of common shares only, ADR's and mutual funds being excluded. We remove penny stocks (consistently trading below \$3), micro-caps (consistently capitalized below \$100M), stocks with consistent extreme returns, and companies for which insufficient return and market capitalization data is available (more than 10% missing). For the purpose of this study, the exclusion criteria are based on the eleven-year window to determine the universe.

As explained in Section 2.4.1, at each portfolio reconstruction date we form factor-mimicking portfolios; see Figure 3. The reconstruction frequency is the frequency at which we form these portfolios. At these dates, the constituents are selected from the universe by looking at certain stock-specific attributes. If so-called breakpoints are used, the companies are ordered with respect to the attribute, and all stocks with a rank falling in the interval defined by the breakpoints will be taken into the factor-mimicking portfolio.

Figure 3

Construction of factor-mimicking portfolios



The factor-mimicking portfolios are value weighted. Between two reconstruction dates, the portfolio constituents and the number of shares invested in each stock are kept fixed, with the exception of companies becoming inactive, for instance due to acquisitions or bankruptcies. In this case, we assume that the stock is sold at the last date with a valid quote in the Compustat database, and that the proceeds are immediately invested into the remaining constituents, proportionally to the stock weights on the transaction day. Dividends and other cash payments of a stock are immediately reinvested in the same stock, so that we work with total returns; see Section 2.

Note that even though for each stock the number of held shares is kept fixed during the reconstruction period, the portfolio will not stay value weighted due to the dividends. For this reason, the portfolio is rebalanced at predefined rebalancing dates. On a rebalancing day, the necessary trades are performed to make the portfolio value-weighted again.

At the end of the return periods, each stock weight changes due to the inhomogeneous performance of stocks. Between rebalancing dates and provided no stock leaves the portfolio, weights evolve according to

$$w_i(t) = \frac{w_i(t-1)(1+r_i(t))}{1+\sum_{k=1}^{K} w_k(t-1)r_k(t)}\,. \tag{24}$$

3.1.2 Factor construction parameters

The factors that will be used in our study are summarized in Table 3, with corresponding attributes, annualized volatilities, and mean returns. For all construction parameters (attributes, breakpoints, reconstruction and rebalancing frequencies), our main worry is to keep factors as simple as possible in order to facilitate their interpretation.

Table 3

Factor description for the 1998–2005 period

Factor	Attribute	Mean	Volatility
Market	-	6.24	16.25
Size (SMB)	Market cap	13.13	16.4
Value/Growth (HML)	Price to book	6.99	12.51
Dividend	Dividend yield	-6.14	23.88
Liquidity	Nb Traded/Nb Outstanding	3.89	24.35
Momentum	Past 6 month returns	3.79	22.88
Volatility	Realized volatility	0.77	13.32
Energy	GICS Code	13.33	21.33
Materials	GICS Code	8.03	20.78
Industrials	GICS Code	7.49	17.04
Consumer Discretionary	GICS Code	6.05	19.46
Consumer Staples	GICS Code	4.04	13.83
Health Care	GICS Code	7.2	14.66
Financials	GICS Code	8.2	18.48
Information Technology	GICS Code	4.77	35.16
Telecommunication Services	GICS Code	-0.66	23.98
Utilities	GICS Code	6.51	16.8

Recall that we have introduced two versions of Fama-French type factors in Section 2.4.1: a simplified version constructed from a single attribute only, and the originally proposed factors based on two attributes, as given by (17) and (18). Moreover, for certain factors the literature proposes a variety of possible attributes. For instance, value/growth (HML) could also be calculated by using price-to-earnings instead of price-to-book ratio. We have seen four different versions of momentum factors. Since versions of the same type are highly correlated, we wanted to make an optimal choice among them. To this end, we compared the mean sample R^2 of test portfolios[13] for all possible combination of factors. We found

[13]We introduce this notion in Section 3.2.

Table 4
Mean sample R^2 of 900 test portfolios

Models	1996-2001	2001-2005	Total
Maximum R^2	0.84	0.91	1.75
RiskMetrics selected factors R^2	0.83	0.90	1.73
Minimum R^2	0.80	0.88	1.68

that the version of the factor has only marginal impact on the mean sample R^2, as shown in Table 4. For this reason, we decided to work with the simplified versions only. For HML, we took the price-to-book ratio (lagged by six months) and the momentum factor was based on the past six months performance.

The Compustat company level mnemonics we have use for the attributes are *prccm* and *trfm* for respectively the monthly pricing and multiplication factor for calculating the total return, *mkvalm* for the market capitalization based on actively traded issues, *mkbk* for the price-to-book ratio, *dvpsxm* for cash dividends per share, *cshtrm* for monthly number of traded shares, and *cshoq* for the number of outstanding shares (including non-actively traded issue).

Regarding the breakpoints, the factors will use either no breakpoints (market, industry sectors), a 0.5 breakpoint for SMB or the $\{0.3, 0.7\}$ breakpoints for the remaining style factors (HML, dividend, momentum, volatility, and liquidity).

Reconstruction and rebalancing frequencies are set to annual and semi-annual respectively, except for the momentum factor, where every month, portfolios are reconstructed based on the past six months' winners and losers. In this case, reconstruction and rebalancing are monthly.

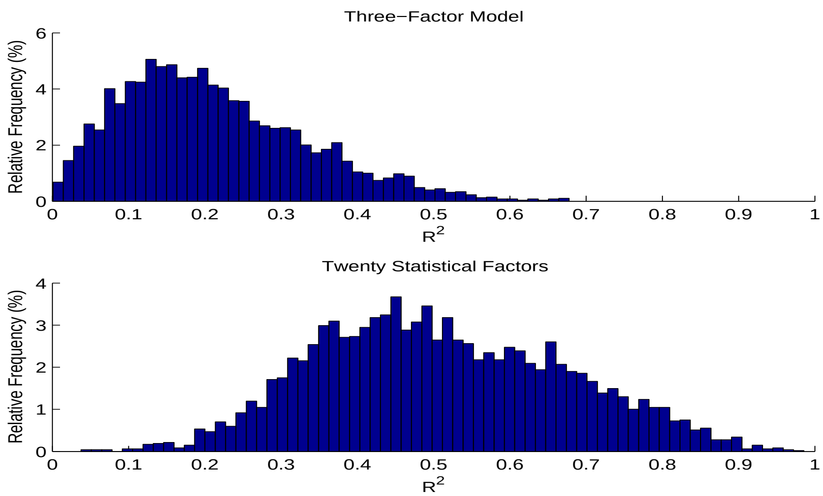
3.2 Test portfolio framework

As a first step in assessing the factors of Table 3 we look at their explanatory power against single stocks. For the 1996–2006 period, we regress the equities on the market factor, SMB, and HML (three-factor model) and against the first twenty statistical factors extracted using asymptotic PCA. The universe contains 4685 US equities. Results are shown in Figure 4. The median R^2 of 0.19 for the three-factor model is fairly low. This could mean three things:

Figure 4

4685 US stocks regressed individually against the three-factor model and 20 statistical factors

The median R^2 are 0.19 and 0.49 for the three-factor model and the twenty statistical factors, respectively.



1. We have not used sufficiently many factors,
2. The specific component dominates in case of individual stocks returns, resulting in a lot of noise, or
3. The factors we use are not well constructed.

The first point concerns the question of the number of factors. Empirical studies (Bai and Ng 2002; Connor and Korajczyk 1993) have shown that the optimal number of factors in the cross-section of NYSE and AMEX stock exchanges is below ten. This suggests that although adding factors will increase R^2, it will diminish the accuracy of our model, and that this route should not be followed.

Regarding point two, the median R^2 of the twenty statistical factors confirms that, unsurprisingly, stocks carry a large specific risk. Even with the twenty statistical factors—which by construction explain the variance optimally—the median sample R^2 does not exceed 50%. At the same time, this also indicates that we cannot verify hypothesis number three using single stocks. Because factors seek to account for the systematic components, it would be counterproductive to assess the quality of factors on returns dominated by specific risk. Instead we should test our factors on portfolios, where the specific part is

Table 5
Number of securities per style on portfolio formation dates

	31 December 1995			**31 December 2000**		
	Value	Neutral	Growth	Value	Neutral	Growth
Large	180	236	222	148	188	274
Medium	222	204	201	198	219	189
Small	229	190	208	261	199	144

largely diversified away.[14] This led us to develop our test portfolio framework.

We form value-weighted random test portfolios containing 50 stocks each. On 31 December 1995 and 31 December 2000, we form nine "style boxes" according to market capitalization and price-to-book ratio ranks. For both attributes, we use the breakpoints $\{0.33, 0.66\}$. In 1995, we remove stocks with market capitalization below \$250M, and in 2000, remove those below \$400M. For each style box, we form 100 portfolios of 50 stocks chosen at random with probability proportional to the stock's market capitalization on that day. In Table 5, we display the number of stocks per style on the portfolio formation dates. The ad hoc additional limits on market capitalization (respectively \$250M and \$400M) were set to obtain a similar amount of stocks and a balanced distribution within each style.

The constituents do not change between formation dates (except if companies become inactive), and the portfolios are rebalanced annually. In Figure 5, we have represented the portfolios as of December 1995, ranked according to both average size and average price-to-book. The ranking is done with respect to the 1892 stocks in the universe and uses the average size/price-to-book of each portfolio.

The clustering of the portfolios in each style is a result of the averaging. The portfolio dispersion along the size dimension is thin since the large cap stocks tend to dominate in value-weighted portfolios. The distribution of portfolios along the price-to-book dimension is more widespread. Note that a couple of extreme price-to-book values stocks lead to points which lie at the far right edge of the box.

In Table 6, we provide average annual performance and volatilities for each style in both periods. Note the spectacular fall in 2001–2005 of all large cap and growth portfolio average returns. Alone, the small value segment improved substantially compared to the first period.

[14] This point has also been stressed by MacQueen (2003) .

Figure 5

Location of portfolios in style box on 31 December 1995

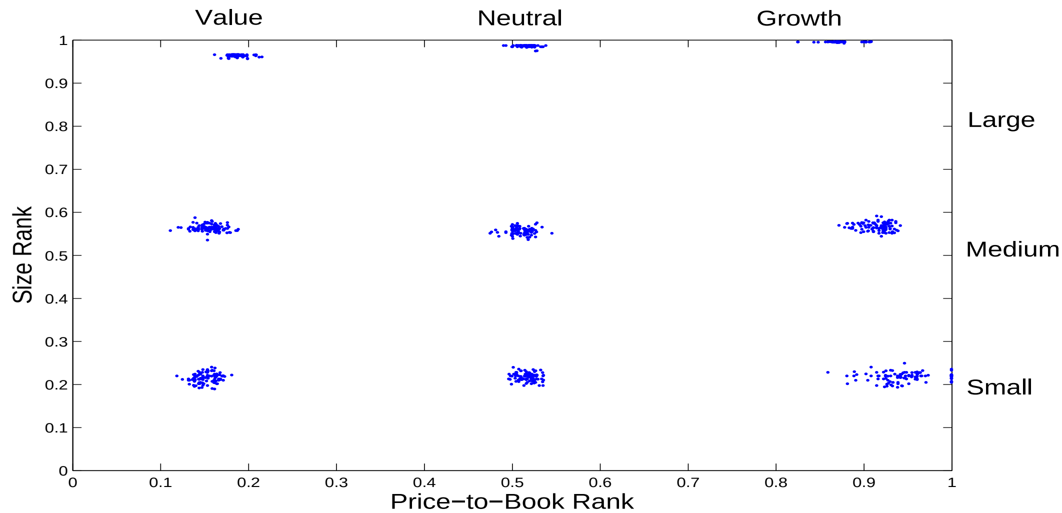


Table 6

Portfolio average annual performance and volatility, 1996–2005

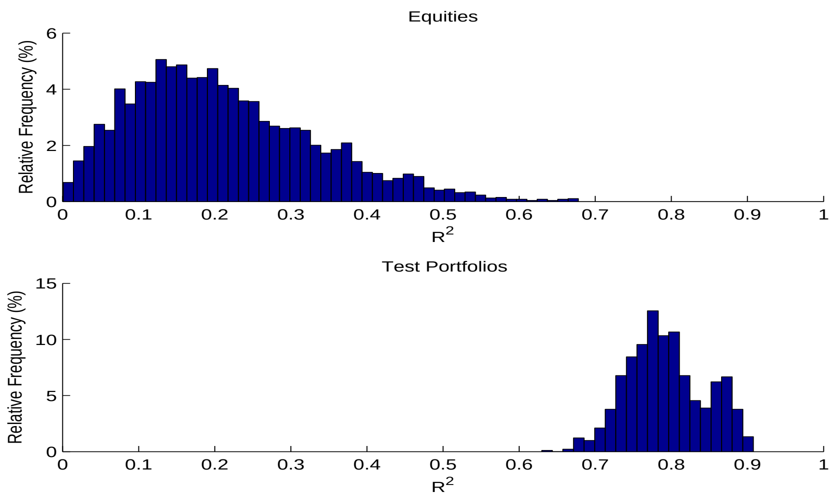
1996-2000:	**Performance**			**Volatility**		
	Value	Neutral	Growth	Value	Neutral	Growth
Large	14.59	20.42	21.50	20.81	16.67	18.55
Medium	14.14	13.72	7.12	20.62	22.04	25.70
Small	13.94	15.33	20.15	18.4	24.96	35.17

2001-2005:	**Performance**			**Volatility**		
	Value	Neutral	Growth	Value	Neutral	Growth
Large	3.6	3.19	-3.40	16.65	15.31	16.62
Medium	14.93	10.48	4.23	18.69	16.81	23.55
Small	15.43	10.17	3.21	19.77	20.36	23.83

Figure 6

Individual stocks and test portfolios regressed against the three-factor model, 1996–2005

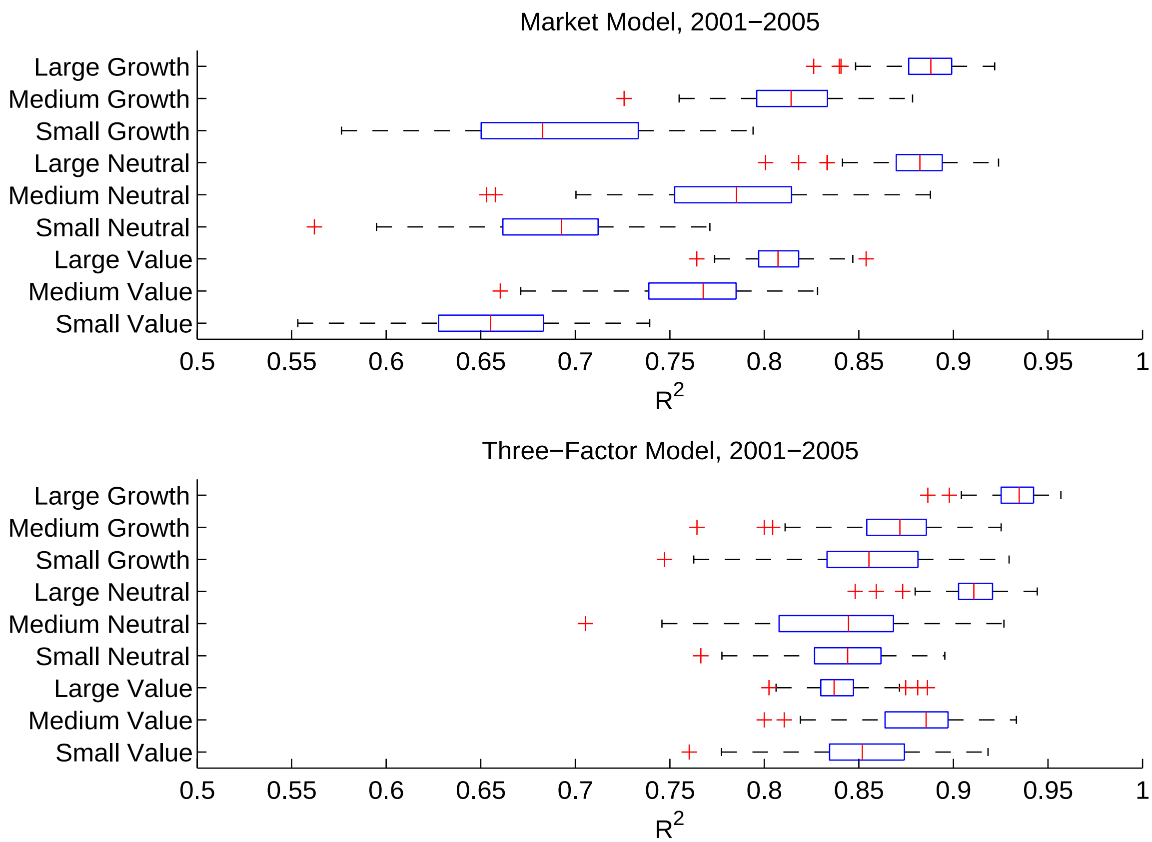
The median sample R^2 are 0.19 for the stocks and 0.79 for the test portfolios.



Now that we have set the framework, we return to our initial task—to assess our factors. We began by regressing the three-factor model on US stocks and concluded that the specific component was overwhelming. Let us now compare between having stocks or portfolios as the testing environment. In Figure 6 we plot the same histogram as before (US stocks regressed on the three-factor model) and add the sample R^2 distribution of the 900 portfolios regressed on the three-factor model. The median sample R^2 jumps from 0.19 to 0.79, indicating the presence of a large systematic component in the portfolio returns captured by the factors. Since assessing factors is also measuring how well they capture this systematic part, we adopt the test portfolio framework for factor assessment.

Figure 7

Comparison of the market and three-factor models applied to the test portfolios



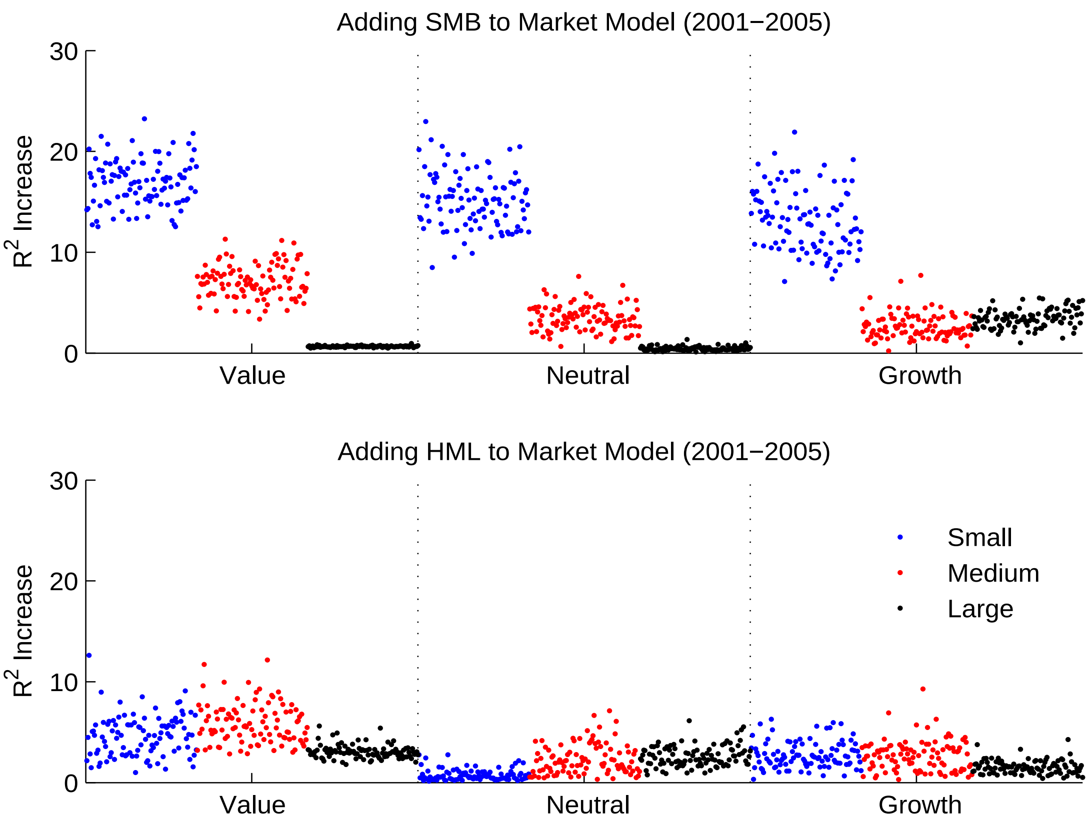
3.3 Assessing the explanatory power of factors

We have established that it is beneficial to analyze factors with respect to test portfolios. In what follows, we examine the R^2 for test portfolios under various models and during different time periods.

As a first check, we analyze the impact on R^2 when the two Fama-French factors SMB and HML are added to the market model, resulting in the so-called three-factor model. The results are summarized in Figure 7. For each portfolio, monthly returns in the period 31 January 2001 to 31 December 2005 are regressed against the market factor (top graph) and the factors of the three-factor model (bottom graph).

Figure 8

Absolute change in R^2 when SMB (top) or HML (bottom) is added to the market factor



For every investment style, the corresponding R^2 are displayed by means of a boxplot.[15] It is striking that including SMB and HML leads to a significant increase in R^2, particularly for all of the small cap portfolios. Also the dispersion of the R^2 is visibly reduced. The large portfolios' R^2 are high already in the market model, and less improved by inclusion of the Fama-French factors. Summarizing, it appears that SMB and HML add a lot of explanatory power to the market model, indicating the market model misses one or more common components. Formal statistical tests could be conducted, but we believe the visual results are conclusive enough.

[15] The first quartile $q_{0.25}$ and third quartile $q_{0.75}$ of the data bound the box. The box is divided by the median. Observations more than 1.5 times the interquartile range $q_{0.75} - q_{0.25}$ away from the edges of the box are considered as outliers and marked red. The whiskers show the farthest points which are not outliers.

Table 7
Definition of factor models

Model					**Factors**			
	Market	SMB	HML	Momentum	Dividend	Liquidity	Volatility	Sectors
Model 1	xx							
Model 2	xx	xx						
Model 3	xx		xx					
Model 4	xx	xx	xx					
Model 5	xx	xx	xx	xx				
Model 6	xx	xx	xx	xx	xx			
Model 7	xx	xx	xx	xx	xx	xx		
Model 8	xx	xx	xx	xx	xx		xx	
Model 9	xx	xx	xx	xx	xx	xx	xx	
Model 10	xx	xx	xx	xx	xx	xx	xx	xx
Model 11	xx							xx

Next, we compare SMB and HML in terms of added R^2 for each of the nine investment styles. In Figure 8, we plot the absolute changes of R^2 when the Fama-French factors are added to the market model. It is apparent that during the years 2001–2005, the SMB factor added more than the HML. Not surprisingly, SMB is particularly effective for small caps and HML for value stocks.

To get a feeling for achievable R^2 values in the test portfolios, we have also looked at other factor combinations. Table 7 defines these combinations; Table 8 shows the numerical results. There are two interesting observations. First, in the 1996–2000 period, the observed R^2 are lower than during 2001–2005. A reason might be the burst of the technology bubble and the bear market in 2001–2003, both of which fell in the second period, but this would need to be investigated further.

Second, in general, the R^2 is increased when factors are added to a certain model. However, more factors do not necessarily entail a higher R^2: the market together with SMB, HML, the momentum and dividend factor (Model 6) has apparently a higher explanatory power than the market combined with the ten sector factors (Model 11). Moreover, Model 10 shows relatively high interquartile range for the R^2; this might

Table 8
R^2 of test portfolios

Model	**1996-2001**		**2001-2005**	
	Median R^2	Interquartile range	Median R^2	Interquartile range
Model 1	0.58	0.24	0.78	0.14
Model 2	0.70	0.17	0.84	0.06
Model 3	0.72	0.16	0.82	0.12
Model 4	0.78	0.09	0.86	0.07
Model 5	0.79	0.09	0.88	0.05
Model 6	0.81	0.08	0.89	0.05
Model 7	0.82	0.09	0.89	0.05
Model 8	0.82	0.08	0.89	0.04
Model 9	0.83	0.08	0.90	0.04
Model 10	0.91	0.08	0.94	0.03
Model 11	0.80	0.16	0.88	0.08

indicate that the sector factors miss something.

Note that we did not test all models that could be formed through combining the factors. Altogether there are seventeen factors available, giving $2^{17} - 1 = 131071$ possible models. This lies at the edge of what can be analyzed with standard computing power.

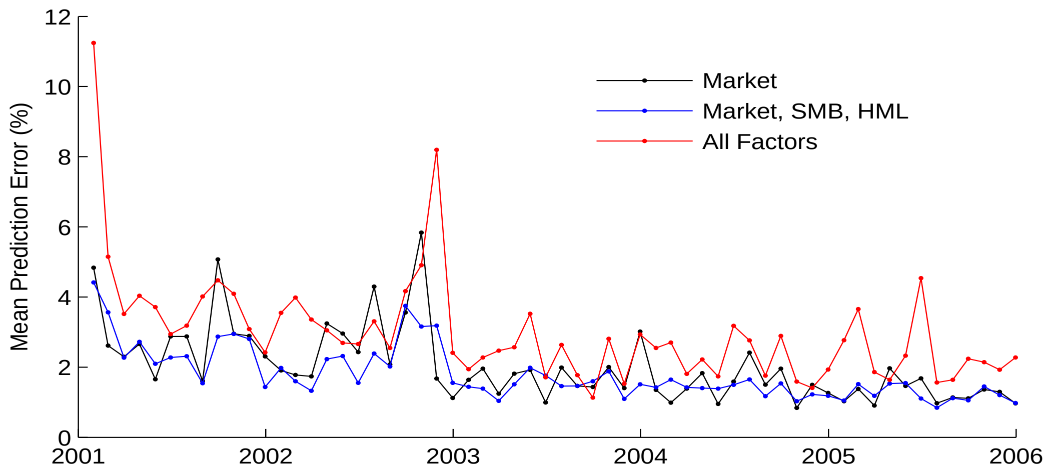
3.4 Prediction error

As we mentioned at the beginning of this section, there is a latent danger of overfitting when there is a one-sided focus on sample R^2. It is always possible to achieve a sample R^2 of one if we add sufficiently many explanatory variables to the model; they do not even have to be related to the returns. In this way, we arrive at a model which explains historical data perfectly, but is most likely extremely poor in predicting. When too many irrelevant variables—variables with a true regression coefficient of zero—are included in a model, the estimation errors of the coefficients distort the predictions.

Figure 9

Prediction error for the test portfolios in the market model, three-factor model, and full model

The EWMA parameter is set to 1.0 and regression coefficients are determined using moving windows of 24 monthly returns.



We illustrate these statements by means of a practical example. We consider the mean prediction error of the test portfolios, which is defined as follows. Assume we estimate the parameters $\alpha_i(t)$ and $\beta_{i\cdot}(t)$ through (13), OLS, or any other estimator. Then the canonical one step ahead predictor of $r_i(t+1)$ is given by

$$\widehat{r}_i(t+1) = \widehat{\alpha}_i(t) + \sum_{j=1}^{p} \widehat{\beta}_{ij}(t)\, f_j(t+1). \tag{25}$$

In other words, we suppose that the factor values $f_j(t+1)$ are known, and predict the next return using the last estimated regression. This is because we want to avoid errors stemming from the prediction of $f_j(t+1)$; we are only interested in the quality of the linear factor model. We calculate the mean error of this predictor across all equities (or portfolios), giving

$$\widehat{PE}(t) = \frac{1}{K} \sum_{i=1}^{K} |\widehat{r}_i(t) - r_i(t)|. \tag{26}$$

Figure 9 graphs the mean prediction error for the 900 test portfolios in the market model, the three-factor model and a model containing all seventeen factors of Table 3.

If all seventeen factors are used, the mean prediction error is highest, Even the simple market model outperforms the seventeen-factor model in terms of mean prediction error. The three-factor model predicts best. Its average mean prediction error during the period 2001-2005 is 1.87, compared to 2.02 for the market model and 2.93 for the seventeen-factor model. Note that the plots of the mean prediction error are all slightly downward sloping. This is caused by the fact that the level of residual volatility was decreasing during the analyzed period.

The results of our experiment are very instructive and indicative of the saying that “less is more”. We mentioned earlier in this paper that it is left to the analysts taste which type of factors he wants to use. Say, a manager wants to employ industry factors. We mention that the GICS classification contains 67 industries, so that there are 67 industry factors in total. When estimating the 67 loadings from only 24 monthly returns, the coefficients are not unique since the design matrix cannot have full rank. When relying on sufficiently many daily returns and factors, this particular problem would most likely disappear. But even then an extremely low predictive power has to be expected for the resulting model.

Statisticians have dealt for a long time with the kind of problems described above. They have developed methods to overcome them, and these are presented next.

3.5 Statistical factor selection

Continuing with the example of industry factors from above, we have seen that the cost of regressing onto all of them can be high. Nevertheless one would like to give the factor modeler the flexibility to use the industry factor type. A natural remedy is to statistically select from the industry factors. The main idea is to regress only on those factors which have explanatory power for the returns and to discard the others.

Statistical variable selection is a field on its own, and it is beyond the scope of this paper to give a complete picture. There is a variety of variable selection techniques, each having its strengths and weaknesses. For this paper we would like to confine ourselves to stepwise procedures, or more specifically, to the so-called forward stepwise selection procedure (FSSP). This is a widely applied and well-studied method, which is moreover intuitive and relatively easy to explain. In what follows, we refer to standard textbooks for the formulas and computational details.

The FSSP starts with no factors in the model and a candidate pool of potentially interesting factors; factors are selected from the pool through an iterative process. In each step of the process, the relevance of each factor is assessed by its p-value in a test of the null hypothesis that the factor’s regression coefficient is zero. Typically, there are two cutoff levels for the p-value, $p_{in} < p_{out}$. A variable can be selected only if its p-value is smaller than p_{in}, and eliminated if its p-value is larger than p_{out}. Common values are $p_{in} = 0.05$ and $p_{out} = 0.10$.

In the first step of the process, the factor with the lowest p-value (as long as this is less than p_{in}) is entered into the model. In the second step, the previously chosen factor is left fixed, and among the remaining factors in the pool, the one is selected that has lowest p-value in the combined model. Then, in addition to what has been done in the first step, one checks whether the first chosen factor in the combined model now has a p-value greater than p_{out}, in which case it is eliminated.[16] One continues this procedure: select the next most significant variable and then eliminate the factor chosen in earlier steps with the highest p-value, provided the latter is larger than p_{out}. The iteration stops when no more factors can be included anymore or when a pre-specified number of iteration steps has been reached.

As an alternative to the models with fixed factors in Figure 9, we also applied the FSSP to the test portfolios. All seventeen factors were included in the candidate pool of factors. We obtained an average mean prediction error of 2.03, which is an improvement of roughly 50% as compared to the model with all factors. This indicates that statistical variable selection should be employed in an equity factor model, most notably when there are many potentially interesting factors.

3.6 Importance of equity factors

To conclude our study, we address the question of equity factor importance. Note that importance has to be understood in a statistical sense. It would be daring to state that an equity factor being more important than another means that the same holds true for the underlying economic sources of risk. Furthermore, establishing a strong statistical relationship (that is, high R^2) between returns and factors does not prove any causality. These fundamental limitations of any statistical modeling should never be forgotten.

Moreover, in the multiple regression context, a generally accepted notion of "variable importance" does not exist. One might take the R^2 coming from the simple regressions of returns onto one factor as a measure for importance, but this definition misses in a certain sense the character of multiple regression, where a combination of factors is used to explain returns. A better measure is the so-called marginal R^2, which is the loss in R^2 when a particular factor is removed from the equity factor model (with fixed factors). High values of the marginal R^2 are interpreted as the factor being important. One caveat is that marginal R^2 are non-additive. It is important to notice that marginal R^2 should be given a relative meaning. Indeed, if factors are considerably correlated, it can well be that all marginal R^2 are relatively

[16] Here elimination means that the factor goes back to the pool again. It can be re-selected in later steps.

Table 9

Importance of factors in test portfolios during 2001–2005

Factor	Frequency chosen (%)	Marginal R^2 (%)
Market	53.6%	37.5%
Size (SMB)	32.2%	8.6%
Value/Growth (HML)	19.5%	6.7%
Dividend	14.1%	17.9%
Liquidity	20.1%	17.9%
Momentum	17.9%	7.0%
Volatility	7.2%	5.5%
Energy	11.8%	6.3%
Materials	12.0%	6.8%
Industrials	10.0%	9.5%
Consumer Discretionary	29.3%	19.4%
Consumer Staples	14.4%	4.6%
Health Care	10.3%	4.2%
Financials	21.4%	12.5%
Information Technology	12.7%	11.4%
Telecommunication Services	9.4%	4.5%
Utilities	15.5%	6.3%

small, even though the R^2 achieved by including all factors is high.

Connor (1995) employed the marginal R^2 to assess factor importance and to compare equity factor models. We would however like to employ the FSSP, and so we slightly modify the definition of marginal R^2. The definition is tailored to our test portfolio framework. For each portfolio and time t we apply FSSP, and in the resulting model we calculate the marginal R^2 for each selected factor. For a specific factor, we take the mean marginal R^2 across all portfolios (or stocks) and times where that factor was chosen. The results are summarized in Table 9.

Not surprisingly, the market factor is clearly most important: it is chosen in 53.6% of the cases, and its average marginal R^2 is 38%. Among the Fama-French type factors, the size factor is chosen most often. For our test portfolios, SMB is more important than HML. The dividend and liquidity factors have a rather considerable mean marginal R^2 when they are chosen. From verifications we have conducted, it appears that the reason for this is that the FSSP typically chooses one of them instead of the market, so that in certain cases they serve as a "placeholder" for the market factor. Concerning the sectors, the consumer discretionary factor seems to be most important, followed by the financial factor.

4 Conclusions

We have given a survey on equity factor models and have compared them with granular models. Our point of view is that equity factor models are primarily needed for risk decomposition, exposure measurement, and for the estimation of covariance matrices used in portfolio optimization. In terms of total risk numbers, the accuracy of equity factor and granular models is comparable. A combined approach, where risk reports are generated simultaneously using an equity factor model and a granular model, opens up new perspectives for risk measurement and management.

Concerning equity factor models, we propose to take predefined factors along with time series regression for the determination of the loadings. The estimated residuals may be further modeled using PCA. We believe that this approach is most flexible. For equity factor modeling, flexibility and customizability are key to success.

In our empirical study, we have evaluated factors proposed by the literature. We mainly focused on factors constructed by the factor-mimicking portfolio approach. We argued that factors should be exposed to real-world applications, and for this reason we introduced a test portfolio framework, which we then used to assess our factors.

An important insight of our study is the fact that there is nothing like a set of "true" factors. It is the views and focus of the analyst which determines the set of equity factors one should use. To avoid overfitting, statistical factor selection methods are recommended.

References

Ang, A. and J. Chen (2005). CAPM over the long run: 1926–2001. Working Paper.

Bai, J. and S. Ng (2002). Determining the number of factors in approximate factor models. *Econometrica 70*(1), 191–221.

Banz, R. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics 9*, 3–18.

Benson, P. and P. Zangari (1997). A general approach to calculating VaR without volatilities and correlations. *RiskMetrics Monitor*, 19–24.

Chan, L., N. Jegadeesh, and J. Lakonishok (1996). Momentum strategies. *Journal of Finance 51*(2), 1681–1713.

Chan, L., J. Karcoski, and J. Lakonishok (1998). The risk and returns from factors. *Journal of Financial and Quantitative Analysis 33*(2), 159–188.

Chen, N. and J. Ingersoll (1983). Exact pricing in linear factor models with finitely many assets: a note. *Jornal of Finance 38*(3), 985–988.

Chen, N., R. Roll, and S. Ross (1986). Economic forces and the stock market. *Journal of Business 59*(3), 383–403.

Connor, G. (1995). The three types of factor models: a comparison of their explanatory power. *Financial Analysts Journal* (3), 42–46.

Connor, G. and R. Korajczyk (1993). A test for the number of factors in an approximate factor model. *The Journal of Finance 48*(4), 1263–1291.

Connor, G. and R. Korajzcyk (1986). Performance measurement with the arbitrage pricing theory: a new framework for analysis. *Journal of Financial Economics* (15), 373–394.

Fama, E. and K. French (1993). Common risk factors in the returns of stocks and bonds. *Journal of Financial Economics 33*, 3–56.

Fan, J., Y. Fan, and J. Lv (2006). Large dimensionsal covariance matrix estimation using a factor model. Working paper.

MacQueen, J. (2003). The structure of multifactor equity models. *Journal of Asset Management 3*(4), 313–322.

Markowitz, H. (1952). Portfolio selection. *Journal of Finance 7*(1), 77–91.

Mina, J. and J. Xiao (2001). *Return to RiskMetrics: The Evolution of a Standard*. RiskMetrics Group.

Northfield (2006). *Everything Everywhere—The Global Equity & Fixed Income Risk Model*. Northfield Information Services, Inc.

Ross, S. (1976). The arbitrage theory of asset pricing. *Journal of Economic Theory 13*, 341–360.

Sharpe, W. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance 19*(3), 425–442.

Stephan, T., R. Maurer, and M. Dürr (2000). A multiple factor model for European stocks. Working paper.

Swinkels, L. and P. van der Sluis (2006). Return-based style analysis with time-varying exposures. *European Journal of Finance*, 529–552.

Tsay, R. (2005). *Analysis of Financial Time Series* (2nd ed.). John Wiley & Sons.

Zangari, P. (2003). Equity risk factor models. In B. Litterman (Ed.), *Modern Investment Management*. John Wiley & Sons.