Article

Custom v. Standardized Risk Models

Zura Kakushadze [1,2,*] and Jim Kyung-Soo Liew [3]

[1] Quantigic® Solutions LLC, 1127 High Ridge Road #135, Stamford, CT 06905, USA;

[2] Business School & School of Physics, Free University of Tbilisi, 240, David Agmashenebeli Alley, Tbilisi 0159, Georgia

[3] The Johns Hopkins Carey Business School, 100 International Drive, Baltimore, MD 21202, USA; E-Mail: kliew1@jhu.edu

* Author to whom correspondence should be addressed; E-Mail: zura@quantigic.com; Tel.: +1-646-221-0440; Fax: +1-646-792-3264.

Abstract: We discuss when and why custom multi-factor risk models are warranted and give source code for computing some risk factors. Pension/mutual funds do not require customization but standardization. However, using standardized risk models in quant trading with much shorter holding horizons is suboptimal: (1) longer horizon risk factors (value, growth, etc.) increase noise trades and trading costs; (2) arbitrary risk factors can neutralize alpha; (3) "standardized" industries are artificial and insufficiently granular; (4) normalization of style risk factors is lost for the trading universe; (5) diversifying risk models lowers P&L correlations, reduces turnover and market impact, and increases capacity. We discuss various aspects of custom risk model building.

Keywords: risk model; multi-factor; risk factor; short horizon; quant trading; style; industry; specific risk; factor risk; portfolio optimization

1. Introduction

In most incarnations, multi-factor risk models for stocks (RM) are based on style and industry risk factors. Industry factors are based on a similarity criterion, stocks' membership in industries under a given classification (e.g., GICS, ICB, BICS, etc.). Style factors are based on some estimated

(or measured) properties of stocks. Examples of style factors are size [1], value and growth [2–8], momentum [9,10], liquidity [11–14], volatility [15], *etc.*[1]

Most commercial RM are standardized (SRM). Majority of their users are institutions with longer term holdings (mutual funds, pension funds, *etc.*). To these users it is important that RM: (i) contain fundamental longer horizon style risk factors (value, growth, *etc.*), and (ii) be standardized, so their risk reporting and cross-institutional communications can be uniform. If a mutual fund risk report says "Portfolio A has exposure B to the risk factor C under the risk model D", its pension-fund client knows what that means. Typically, such portfolios are broad (low turnover allows lower cap/liquidity holdings) and fairly close to SRM coverage, so universe customization is not critical. Standardization is a much higher priority.

Quant trading with shorter-term holding portfolios has opposite priorities. Such strategies trade smaller universes of more liquid stocks (roughly, around 1000–2500 names), and these universes differ from strategy to strategy. Standardization is not critical for typically in-house risk reporting. Customization is much more important for a variety of reasons, which we discuss in more detail below. Here is a summary.

(1) *Short v. Long Horizons.* Strategies with shorter holding periods (e.g., statistical arbitrage and high frequency trading) do not benefit from longer horizon risk factors (value, growth, *etc.*)—shorter-term returns are not highly correlated with quantities such as book value, which updates quarterly and lacks predictive power for holding horizons measuring in days or intraday. Moreover, when such risk factors are included in regression or optimization of shorter-term returns, they add noise to the holdings and trades, thereby increasing trading costs and reducing profitability.

(2) *Inadvertent Alpha Neutralization.* If, by design, some alphas are taking certain risk exposure, optimization using SRM typically will (partially) neutralize such alphas by producing a portfolio with more balanced risk exposure. e.g., if alphas are intentionally skewed toward small cap value, optimization using SRM, which typically contains "size" risk factor, will muddle such alphas by producing a portfolio with more balanced larger cap exposure. In such cases it is important to use custom RM (CRM) with the corresponding undesirable risk factors carefully omitted.

(3) *Insufficient Industry Granularity.* SRM coverage universe U_{SRM} is "squeezed" into a modest number of fixed standardized "industries" (SRMI). Typically this reduces industry granularity, which adversely affects hedging industry risk. Thus, under a given industry classification (e.g., GICS, ICB, BICS, *etc.*), the true number of industries[2] into which a given trading universe U falls can be sizably higher than the number of SRMI. Also, different trading universes U_1 and U_2 fall into different (numbers of) true industries, while in SRM they are classified into the same SRMI.

(4) *Trading v. Coverage Universe.* Restricting U_{SRM} to a substantially smaller trading universe U yields side effects: (i) for certain U one can have empty SRMI (e.g., U contains no telecom

---

[1]   For an additional partial list (with some related literature), see, e.g., [16–48], and references therein. For a literature survey, see, e.g., [49].

[2]   Naming conventions vary by industry classification. "Industry" here refers to the most detailed level (*i.e.*, the terminal branch) in a given classification tree (see Subsection 2.1 for details).

stocks), with no option to omit them from the factor covariance matrix (FCM) computation;[3] and (ii) it spoils normalization of style factors—those normalized across U_{SRM}, e.g., conformed to a (log-)normal distribution.

(5) *Herd Effect.* Overusing SRM makes alphas more correlated, so when a shop blows up, it drags others with it (e.g., Aug'07 Quant Meltdown). This adversely affects shorter holding trading. Longer horizon strategies simply weather the storm.

(6) *Diversifying Risk Models Lowers P&L Correlations.* When evaluating two RM against each other (in regression or optimization of the same returns), one looks at both (A) the relative performance and (B) P&L correlation. If the P&L correlation is low enough, it is more optimal to run a combined strategy using both RM thereby reducing portfolio turnover and market impact, and increasing capacity and P&L.

(7) *Pros and Cons of Custom Risk Models.* CRM further provide: (i) ability to compute RM based on custom universes and add/subtract risk factors; (ii) transparency; (iii) flexibility with update frequency (daily, weekly, *etc.*) and integration into test/production environments. However, the portfolio manager (PM) must compute CRM as opposed to receiving a file from an SRM provider.[4] Data required to build CRM for shorter horizon trading is typically available to PM, albeit self-consistently computing FCM and idiosyncratic risk (ISR) is not common knowledge.

We discuss the above points (1)–(7) in more detail below. Section 2 discusses RM in general, differences in their use in regression and optimization, and point (2). Section 3 discusses decoupling of time horizons relevant to point (1). Section 4 discusses points (3) and (4). We conclude in Section 5 by discussing points (5)–(7). Appendix A contains R code for some style risk factors. Appendix B contains C code for symmetric matrix inversion. Appendix C contains legal disclaimers.

2. Multi-factor Risk Models

In RM, a sample covariance matrix (SCM) C_{ij} for N stocks, $i, j = 1, \ldots, N$ (computed based on time series of stock returns) is modeled by Γ_{ij} given by

$$\Gamma \equiv \Xi + \Omega \ \Phi \ \Omega^T \tag{1}$$

$$\Xi_{ij} \equiv \xi_i^2 \ \delta_{ij} \tag{2}$$

where δ_{ij} is the Kronecker delta; Γ_{ij} is an $N \times N$ matrix; ξ_i is specific a.k.a. idiosyncratic risk (ISR) for each stock; Ω_{iA} is an $N \times K$ factor loadings matrix (FLM); and Φ_{AB} is a $K \times K$ factor covariance

---

[3] In fact, SRM sometimes may use a (small) subset U_* of U_{SRM} to compute FCM—the required historical data may not be available for the entire U_{SRM}. However, it may be (and typically is) available for the trading universe U, so U, not artificial U_*, should be used for computing FCM.

[4] Principal components provide "customization" to some extent. However, see footnote 8.

matrix (FCM), $A, B = 1, \ldots, K$. *i.e.*, the random processes Υ_i corresponding to N stock returns are modeled via N random processes χ_i (ISR) together with K random processes f_A (factor risk):

$$\Upsilon_i = \chi_i + \sum_{A=1}^{K} \Omega_{iA} f_A \tag{3}$$

$$\langle \chi_i, \chi_j \rangle = \Xi_{ij} \tag{4}$$

$$\langle \chi_i, f_A \rangle = 0 \tag{5}$$

$$\langle f_A, f_B \rangle = \Phi_{AB} \tag{6}$$

$$\langle \Upsilon_i, \Upsilon_j \rangle = \Gamma_{ij} \tag{7}$$

The main reason for replacing the sample covariance matrix C_{ij} by Γ_{ij} is that the off-diagonal elements of C_{ij} typically are not expected to be too stable out-of-sample. A constructed factor model covariance matrix Γ_{ij} is expected to be much more stable as the number of risk factors, for which FCM Φ_{AB} needs to be computed, is $K \ll N$. Also, if $M < N$, where $M + 1$ is the number of observations in each time series, then C_{ij} is singular with M nonzero eigenvalues. Assuming all $\xi_i > 0$ and Φ_{AB} is positive-definite, then Γ_{ij} is automatically positive-definite (and invertible).

2.1. Industry Risk Factors

RM can mix risk factors of different types. Typically, the most numerous factors are industry factors. In its simplest incarnation, industry FLM $\Omega_{iA}^{\mathrm{ind}}$, $A = 1, \ldots, K_{\mathrm{ind}}$ is a binary matrix of 1s and 0s indicating whether a stock belongs to a given industry:

$$\Omega_{iA}^{\mathrm{ind}} = \delta_{G(i),A} \tag{8}$$

$$G : \{1, \ldots, N\} \mapsto \{1, \ldots, K_{\mathrm{ind}}\} \tag{9}$$

$$\sum_{i=1}^{N} \Omega_{iA}^{\mathrm{ind}} = N_A \tag{10}$$

$$\sum_{A=1}^{K_{\mathrm{ind}}} \Omega_{iA}^{\mathrm{ind}} = 1 \tag{11}$$

where G is the map between stocks and industries, and $N_A > 0$ is the number of stocks in the industry labeled by A (we are assuming there are no empty industries).

More generally, we can allow each stock to belong to multiple industries (e.g., in the case of conglomerates) with some weights ω_{iA}:

$$\Omega_{iA}^{\mathrm{ind}} = \sum_{B \in G(i)} \omega_{iB} \, \delta_{AB} \tag{12}$$

$$\sum_{B \in G(i)} \omega_{iB} = 1 \tag{13}$$

where $G(i) \subset \{1, \ldots, K_{\mathrm{ind}}\}$ and $|G(i)| \equiv n_i$ need not be 1, albeit typically for most stocks $n_i = 1$, and (for conglomerates) $n_i \neq 1$ is not a large number.

Binary $\Omega_{iA}^{\mathrm{ind}}$ can be constructed from binary classifications such as GICS, ICB, BICS, etc. Naming conventions for levels differ by classification, e.g., sector, sub-sector, industry, sub-industry, etc.

"Industry" here means the most detailed level in the classification tree. If there are too many industries (e.g., with small N_A), one can "prune" the tree by merging (small) industries at higher levels, e.g., in the tree "sector → sub-sector → industry", industries are pruned at the sub-sector level.

2.2. Style Risk Factors

Industry factors are based on a similarity criterion: industry membership. Style risk factors are based on some estimated (or measured) properties of stocks. Examples of style factors are size, liquidity, volatility, momentum, growth, value, etc.

For illustrative purposes, let us discuss some style factors in more detail. In Appendix A we give R code for size, liquidity, intraday volatility and momentum style factors.[5]

• Size is a logarithm of market cap, normalized (conformed to normal distribution)—see R code in Appendix A. Note that ADRs are treated separately.

• Liquidity is a logarithm of the average daily dollar volume (ADDV),[6] normalized similarly to size with ADRs treated separately—see Appendix A.

• Volatility style factor can be based on historical (relevant for longer-term models) or intraday (relevant for shorter-term models—see the next section) volatility. One way to define intraday volatility is to use intraday high and low price—see Appendix A. The log of intraday volatility is normalized similarly to size.

• Momentum can be defined as a normalized average over D trading days of d-trading-day moving-average returns (e.g., $d = 5$, $D = 252$)[7]—see Appendix A.

• Value can be defined via a normalized book-to-price ratio (negative values need to be dealt with). Growth requires earnings data. As discussed below, longer horizon style factors such as value and growth do not add value in shorter horizon strategies as book value and earnings data essentially are updated quarterly.

2.3. Factor Covariance Matrix and Specific Risk

Once FLM Ω_{iA} is defined,[8] FCM Φ_{AB} and ISR ξ_i must be constructed. Straightforwardly computing FCM as a sample[9] covariance matrix of the risk factors f_A is insufficient as ISR ξ_i computed using such sample FCM typically are ill-defined. Algorithms for consistently computing FCM and ISR usually are deemed proprietary.

---

[5] Legal disclaimers regarding this code are included in Appendix C.
[6] E.g., over the past 20 trading days. One may prefer to take, say, last 3 months.
[7] In the 0-th approximation, this is a D-day return. Removing outliers introduces d-dependence.
[8] One can also use the first K_{prin} principal components (PC) of SCM C_{ij} as columns in FLM Ω_{iA}. However, the out-of-sample instability in the off-diagonal elements of SCM is also inherited by PC. Furthermore, if $M < N$, SCM is singular and only M PC are available. It is for these reasons that style and industry factors are more widely used in practical applications.
[9] Without delving into details, out-of-sample stability and singularity of FCM when $M < K$ are issues to consider.

2.4. Use of Factor Models

RM have a variety of uses, which are not limited to passively measuring risk exposures, but also include actively hedging such risks, e.g., by requiring that a portfolio be neutral w.r.t. various risk factors and/or such risk exposures be optimized.[10] Here we outline the usage of RM in regression and optimization.

2.4.1. Regression

Let R_i, $i = 1, \ldots, N$ be the stock expected returns. A weighted regression (without intercept) of R over FLM Ω with weights z_i is given by (in matrix notation):[11]

$$\varepsilon \equiv R - \Omega \, Q^{-1} \, \Omega^T \, Z \, R \tag{14}$$

$$Z \equiv \mathrm{diag}(z_i) \tag{15}$$

$$Q \equiv \Omega^T \, Z \, \Omega \tag{16}$$

$$\widetilde{R} \equiv Z \, \varepsilon \tag{17}$$

Here ε_i are the residuals of the weighted regression. Also, note that the "regressed" returns $\widetilde{R}_i$ are neutral w.r.t. the K risk factors corresponding to Ω_{iA}:

$$\sum_{i=1}^{N} \widetilde{R}_i \, \Omega_{iA} = 0, \quad A = 1, \ldots, K \tag{18}$$

If Ω_{iA} includes the intercept (a unit vector) as one of its columns, then we have

$$\sum_{i=1}^{N} \widetilde{R}_i = 0 \tag{19}$$

i.e., in this case the regressed returns are demeaned. The weights z_i can be chosen to be unit weights, or, e.g., $z_i \equiv 1/\sigma_i^2$, where σ_i is some historical volatility of R_i.

Using $\widetilde{R}_i$, a simple mean-reversion strategy can be constructed via

$$D_i = -\gamma \, \widetilde{R}_i \tag{20}$$

where D_i are the desired dollar holdings and $\gamma > 0$ is fixed via

$$\sum_{i=1}^{N} |D_i| = I \tag{21}$$

where I is the total desired investment level. The portfolio (20) is dollar neutral if we have Equation (19). For this weighted regression all we need is FLM Ω_{iA}—FCM Φ_{AB} is not needed,[12] nor is ISR ξ_i, albeit if the latter are known, they can be used (instead of σ_i) in the regression weights via $z_i \equiv 1/\xi_i^2$ (see [50] for details).

---

[10] See, e.g., [50] for a recent discussion.

[11] This is a cross-sectional regression; in R notations $\varepsilon = \mathrm{residuals}\,(\mathrm{lm}\,(R \sim \Omega - 1, \mathrm{weights} = Z))$.

[12] Rotating Ω_{iA} by an arbitrary $K \times K$ nonsingular matrix U_{AB}, $\Omega \to \Omega \, U$, does not change the regression residuals (14) or the risk neutrality conditions (18).

2.4.2. Optimization

In optimization, rather than requiring strict neutrality w.r.t. the risk factors, one can require that risk exposure be optimized, albeit one can do both (see below). In its simplest incarnation, optimization requires that the Sharpe ratio of the portfolio (using the notations of the previous subsection)

$$S = \frac{\sum_{i=1}^{N} D_i R_i}{\sqrt{\sum_{i,j=1}^{N} C_{ij} D_i D_j}} \to \max \tag{22}$$

Assuming no costs, constraints or bounds, the Sharpe ratio is maximized by

$$D_i = \zeta \sum_{j=1}^{N} C_{ij}^{-1} R_j \tag{23}$$

where ζ is fixed via (21). In the RM context, one replaces C_{ij} via Γ_{ij}, which gives:[13]

$$D_i = \frac{\zeta}{\xi_i^2} \left(R_i - \sum_{j=1}^{N} \frac{R_j}{\xi_j^2} \sum_{A,B=1}^{K} \Omega_{iA}\, \Omega_{jB}\, \widetilde{Q}_{AB}^{-1} \right) \tag{24}$$

$$\widetilde{Q}_{AB} \equiv \Phi_{AB}^{-1} + \sum_{i=1}^{N} \frac{1}{\xi_i^2}\, \Omega_{iA}\, \Omega_{iB} \tag{25}$$

where Φ_{AB}^{-1} is the inverse[14] of Φ_{AB}, and $\widetilde{Q}_{AB}^{-1}$ is the inverse of $\widetilde{Q}_{AB}$. The desired dollar holdings D_i are not neutral w.r.t. the risk factors, nor are they dollar neutral. Dollar and/or various risk factor neutrality can be achieved via optimization with homogeneous linear constraints (see [50]). Note that to compute the desired dollar holdings (24), we need not only FLM Ω_{iA}, but also FCM Φ_{AB} and ISR ξ_i. This is a key difference between using RM in optimization *vs.* regression.

2.4.3. "Risk-taking" Alphas

While in optimization the resulting desired dollar holdings D_i are not neutral w.r.t. the risk factors, they are "approximately" neutral in the sense that the deviation from neutrality is due to ISR. Indeed, from Equation (24) we have

$$\sum_{i=1}^{N} D_i\, \Omega_{iA} = \zeta \sum_{i=1}^{N} \frac{R_i}{\xi_i^2}\, \Omega_{iB}\, \Delta_{BA}^{-1} \tag{26}$$

$$\Delta_{AB} \equiv \delta_{AB} + \sum_{i=1}^{N} \sum_{C=1}^{K} \frac{1}{\xi_i^2}\, \Phi_{AC}\, \Omega_{iC}\, \Omega_{jB} \tag{27}$$

where Δ^{-1} is the inverse of $\Delta = \Phi\, \widetilde{Q}$. Let $\Phi_{AB} \equiv \kappa\, \widehat{\Phi}_{AB}$, and let us take the limit $\kappa \to \infty$ with $\widehat{\Phi}_{AB}$ = fixed. In this limit the factor risk dominates and ISR is negligible, so optimization reduces to the weighted regression of the previous sub-subsection (see [50] for details), and we have

$$\sum_{i=1}^{N} D_i\, \Omega_{iA} \to 0 \tag{28}$$

---

[13] This follows from the expression for the inverse of Γ: $\Gamma^{-1} = \Xi^{-1} - \Xi^{-1}\, \Omega\, \widetilde{Q}^{-1}\, \Omega^T\, \Xi^{-1}$.

[14] Appendix B contains C code for symmetric matrix inversion.

So, regression yields risk neutrality, while optimization produces approximate risk neutrality. Either way, if the returns R_i have exposure to a risk factor, it is either eliminated (regression) or substantially reduced (optimization). This has important implications for using SRM in certain types of trading.

Thus, imagine that *by design* the returns R_i have *desirable* risk exposure, e.g., our strategy could be deliberately skewed toward small cap value stocks, have exposure to momentum, volatility, *etc.*[15] If such a risk factor is included in FLM Ω_{iA}, then using the full FLM in regression and/or optimization would be to the detriment of our alpha. In the regression we can simply omit the corresponding risk factor(s). However, in the optimization merely omitting risk factors will not do—we also need to recompute FCM and ISR anew based on the remaining risk factors, otherwise we will get wrong predictions for the total risk:

$$\Gamma'_{ii} \equiv \xi_i^2 + \sum_{A',B' \in H} \Omega_{iA'} \Omega_{iB'} \, \Phi_{A'B'} \neq \Gamma_{ii} \tag{29}$$

where $H \subset \{1, \ldots, K\}$ is the subset corresponding to the remaining risk factors. In this case SRM simply will not do and CRM is required.

3. Decoupling of Time Horizons (Frequencies)

There is an important fundamental concept, which can be stated as decoupling of time horizons (or, equivalently, frequencies or wavelengths). In a nutshell, what happens at time horizon T_1 (or frequency $f_1 = 1/T_1$) is not affected by what happens at time horizon T_2 (or frequency $f_2 = 1/T_2$) if T_1 and T_2 are vastly different. By time horizon we mean the relevant time scales. e.g., the time horizon for a daily close-to-close return is 1 day. In terms of returns, the decoupling can be restated as the returns for long-term horizons T_1 being essentially uncorrelated with the returns for short-term horizons T_2. Here is a simple argument.

3.1. Short v. Long Time Horizons

Here is a simple argument for a single stock (or security). Consider a time interval from time t_0 to time $t_M > t_0$. Let us divide it into M intervals $t_0, t_1, \ldots, t_{M-1}, t_M$. For simplicity, we can assume that these intervals are uniform, $t_s = t_0 + s \, \Delta t$, $s = 0, \ldots, M$, albeit this is not critical here. Let the stock prices at times $t = t_s$ be $P(t_s)$. Let us define the return from time t to time t' as

$$R(t, t') \equiv \ln\left(\frac{P(t')}{P(t)}\right) \tag{30}$$

Then we have

$$\widetilde{R} \equiv R(t_0, t_M) = \sum_{s=1}^{M} R_s \tag{31}$$

$$R_s \equiv R(t_{s-1}, t_s) \tag{32}$$

---

[15] Real-life alphas often have sizable exposure to risk—a real-life alpha is any reasonable expected return. e.g., momentum strategies often have substantial exposure to risk. Furthermore, there is no "perfect" risk model. Otherwise, there would only be mean-reversion caused by temporary trading imbalances. For a complementary discussion, see, e.g., [35].

Now let us ask the following question. How correlated is the return R_M for the most recent period (*i.e.*, t_{M-1} to t_M) with the return $\widetilde{R}$ for the entire period (*i.e.*, t_0 to t_M)? To define a "correlation", we need multiple observations. So, we hang yet another index onto our returns, call it α, where $\alpha = 1, \ldots, p$ labels different periods t_0^α to t_M^α, each consisting of M periods (t_{s-1}^α to t_s^α, $s = 1, \ldots, M$), and we wish to compute the correlation between the return for the last such period R_M^α and the return for the entire such period $\widetilde{R}^\alpha$, and α (not s) labels the series (which is a time series) over which the correlation is computed. For simplicity, we can assume that the periods labeled by α are "tightly packed", *i.e.*, $t_M^{\alpha-1} = t_{M-1}^\alpha$, albeit this is not crucial here. We then have $p + M$ time points τ_r, $r = 0, 1, \ldots, p + M - 1$ and, consequently, $p + M - 1$ returns $\widehat{R}_r$, $r = 1, \ldots, p + M - 1$, where

$$\tau_r \equiv t_0 + r\,\Delta t, \quad r = 0, 1, \ldots, p + M - 1 \tag{33}$$

$$\widehat{R}_r \equiv \ln\left(\frac{P(\tau_r)}{P(\tau_{r-1})}\right), \quad r = 1, \ldots, p + M - 1 \tag{34}$$

$$R_s^\alpha = \widehat{R}_{s+\alpha-1}, \quad s = 1, \ldots, M, \quad \alpha = 1, \ldots, p \tag{35}$$

$$\widetilde{R}^\alpha = \sum_{r=\alpha}^{M+\alpha-1} \widehat{R}_r \tag{36}$$

Note that p can be much larger than M, in fact, we will assume this to be the case.[16]

With the covariance $\langle *, * \rangle$ defined as above, we have

$$\sigma_s^2 \equiv \langle R_s, R_s \rangle = \frac{1}{p} \sum_{\alpha=1}^{p} \widehat{R}_{s+\alpha-1}^2 - \frac{1}{p^2}\left(\sum_{\alpha=1}^{p} \widehat{R}_{s+\alpha-1}\right)^2 \approx \sigma^2 \tag{37}$$

where

$$\sigma^2 \equiv \frac{1}{p} \sum_{r=1}^{p} \widehat{R}_r^2 - \frac{1}{p^2}\left(\sum_{r=1}^{p} \widehat{R}_r\right)^2 \tag{38}$$

and we have used the fact that $p \gg M$, which implies that all M variances σ_s^2 are approximately the same. We then have

$$\langle R_s, R_{s'} \rangle \equiv \sigma_s\,\sigma_{s'}\,\Psi_{ss'} \approx \sigma^2\,\Psi_{ss'} \tag{39}$$

$$\langle \widetilde{R}, \widetilde{R} \rangle = \sum_{s,s'=1}^{M} \langle R_s, R_{s'} \rangle \approx \sigma^2 \sum_{s,s'=1}^{M} \Psi_{ss'} \tag{40}$$

$$\langle R_s, \widetilde{R} \rangle = \sum_{s'=1}^{M} \langle R_s, R_{s'} \rangle \approx \sigma^2 \sum_{s'=1}^{M} \Psi_{ss'} \tag{41}$$

where $\Psi_{ss'}$ is the $M \times M$ correlation matrix of returns R_s, and $\Psi_{ss} = 1$. We have

$$\rho_s \equiv \mathrm{Cor}(R_s, \widetilde{R}) \approx \frac{\sum_{s'=1}^{M} \Psi_{ss'}}{\sqrt{\sum_{s,s'=1}^{M} \Psi_{ss'}}} = \frac{1 + \psi_s}{\sqrt{M + \sum_{s=1}^{M} \psi_s}} \tag{42}$$

---

[16] To illustrate, if, say, Δt is 1 day, then we are computing the correlation of the M-day moving average return $\overline{R}^\alpha \equiv \frac{1}{M}\widetilde{R}^\alpha$ with the last daily return in the moving average, and we have p rolling periods like this. We have $p + M$ dates and, consequently, $p + M - 1$ daily returns.

where

$$\psi_s \equiv \sum_{s' \in J_s} \Psi_{ss'} \tag{43}$$

and $J_s \equiv \{1, \ldots, M\} \setminus \{s\}$.

Because we have $M \ll p$, the matrix $\Psi_{ss'}$ is approximately "self-similar" in the sense that all $m \times m$ sub-matrices of $\Psi_{ss'}$ with $s, s' \in \{k, k+1, \ldots, k+m-1\}$, $k = 1, \ldots, M-m+1$ and $1 < m < M$ (*i.e.*, for each m there are $M - m + 1$ such sub-matrices) are approximately the same. Put differently, $\Psi_{ss'}$ approximately depend only on the difference $s - s'$, and, in fact, only on $|s - s'|$ since $\Psi_{ss'}$ is symmetric. Let $\Psi_{s,1} \equiv \eta_{s-1}$, $s > 1$. Then we have $\Psi_{ss'} \approx \eta_{|s-s'|}$, $s \neq s'$, and

$$\psi_M \approx \sum_{s=1}^{M-1} \eta_s \tag{44}$$

$$\sum_{s=1}^{M} \psi_s \approx 2 \sum_{s=1}^{M-1} (M - s)\, \eta_s \tag{45}$$

To estimate the correlation ρ_M, we need to make some assumptions about the correlations η_s. A reasonable assumption is that the correlations $\Psi_{ss'}$ decay as $|s - s'|$ grows. e.g., we can assume that $|\Psi_{ss'}| \leq \lambda^{|s-s'|}$ for some positive $\lambda < 1$, or, equivalently, that $\eta_s = \widetilde{\eta}_s \, \lambda^s$, $s = 1, \ldots, M-1$, where $|\widetilde{\eta}_s| \leq 1$. We then have

$$\psi_M \approx f(\lambda) \equiv \sum_{s=1}^{M-1} \widetilde{\eta}_s \, \lambda^s \tag{46}$$

$$\sum_{s=1}^{M} \psi_s \approx 2\left[M \, f(\lambda) - \lambda \, f'(\lambda)\right] \tag{47}$$

$$f'(\lambda) \equiv \frac{\partial f(\lambda)}{\partial \lambda} \tag{48}$$

and our correlation ρ_M reads

$$\rho_M \approx \frac{1 + f(\lambda)}{\sqrt{M\left[1 + 2\, f(\lambda)\right] - 2\lambda \, f'(\lambda)}} \tag{49}$$

We also have

$$|f(\lambda)| \leq \frac{\lambda - \lambda^M}{1 - \lambda} \approx \frac{\lambda}{1 - \lambda} \tag{50}$$

$$|f'(\lambda)| \leq \frac{1 + (M - 1)\, \lambda^M - M\, \lambda^{M-1}}{(1 - \lambda)^2} \approx \frac{1}{(1 - \lambda)^2} \tag{51}$$

where we have taken into account that $M \gg 1$.

So, if $f(\lambda) \geq 0$, we have (assuming M is large)

$$0 < \rho_M \lesssim \frac{1}{(1 - \lambda)\sqrt{M}} \ll 1 \tag{52}$$

and for positive $f(\lambda)$ the bound is even tighter.

What about $f(\lambda) < 0$? First, note that ρ_M is still positive—for it to become negative, the argument of the square root in the denominator in Equation (49) would become negative, which is not possible (see below). However, for negative $f(\lambda)$ *a priori* it might appear that ρ_M need not be small as the denominator in Equation (49) could become small when $f(\lambda) = -1/2 + \epsilon$, where $0 < \epsilon \sim 1/M$. Nonetheless, this cannot be the case if there is randomness in the returns R_s. Indeed, we have[17]

$$\langle \widetilde{R}, \widetilde{R} \rangle \approx \sigma^2 \sum_{s,s'=1}^{M} \Psi_{ss'} \approx \sigma^2 \left(M \left[1 + 2 \, f(\lambda) \right] - 2\lambda \, f'(\lambda) \right) \tag{53}$$

So, the argument of the square root in the denominator of Equation (49) is (up to σ^2) the variance of the return $\widetilde{R}$ for the period $t_M - t_0 = M \, \Delta t$. If there is randomness in the returns R_s, the variance $\langle \widetilde{R}, \widetilde{R} \rangle$ should scale linearly with $t_M - t_0$ and, consequently, with M. If this variance were of order σ^2, this would imply that the returns R_s are highly anti-correlated with each other and the entire process is highly deterministic. Put differently, there would be essentially no dispersion in this case. Under normal circumstances, where we have randomness in the returns R_s, the variance $\langle \widetilde{R}, \widetilde{R} \rangle$ should be of order $M \, \sigma^2$. If there are any negative correlations η_s, they are offset by other positive correlations so that $\langle \widetilde{R}, \widetilde{R} \rangle \sim M \, \sigma^2$ and we have (52).

The upshot is—this is a generalization of our example above—that quantities with long time horizons have low correlations with quantities with short horizons. What happens, say, at milliseconds gets diluted by the time one gets to, say, month-long horizons—and this dilution is due to the cumulative effect of everything that transpires in between such vastly different time scales. Randomness plays a crucial role in this dilution. If things were deterministic, such dilution would not occur.[18]

3.2. Implication for Risk Factors

A practical implication of the above discussion is that care is needed in choosing which risk factors to use in RM depending on what the time horizons of the strategies are for which RM is used. If these horizons are short, then risk factors such as value and growth, whose underlying fundamental data updates quarterly, should not be used as they add no value in short holding (a few days, overnight, intraday, *etc.*) strategies. Here is a simple argument. Consider high frequency trading at, say, millisecond time scales. Does book value make a difference to such trading? The answer is no. What is relevant here is the market microstructure at the millisecond timescales (bid, ask, bid and ask sizes, order book depth, hidden liquidity, posting orders fast on different exchanges, whether the trader's collocation is close to

---

[17] Note that $\rho_M > 0$ unless $f(\lambda) \leq -1$, for which $\langle \widetilde{R}, \widetilde{R} \rangle$ would be negative considering $M \gg 1$.

[18] This is "analogous" to what happens in quantum mechanics and quantum field theory. We put the adjective "analogous" in quotation marks because a stochastic process described by Brownian motion is nothing but Euclidean quantum mechanical particle, so the "analogy" is in fact precise.

the exchange connectivity hub, *etc.*).[19] Whether the book value for stock XYZ is $100M or $1B does not directly affect the market microstructure at millisecond time scales.[20]

On the other hand, quantities such as liquidity and market cap[21] do affect market microstructure. e.g., liquidity affects typical bid/ask sizes, print sizes, etc. More precisely, liquidity computed based on, say, 20-trading-day ADDV indirectly relates to such "micro" quantities because of the expected *linear scaling* of volumes.[22] *i.e.*, even though ADDV is computed using longer horizons, it is a relevant risk factor for shorter horizon strategies precisely because of the aforementioned linear scaling of volumes, allowing an *extrapolation* from longer to shorter horizons.

Similarly, volatility is a relevant style factor. Typically, it is computed as historical volatility of, say, close-to-close returns. As an *extrapolation*—based on the *assumption* that historically more volatile stocks are also more volatile intraday—one can use this style factor for shorter horizon strategies. Preferably, one can also define volatility style factor based on shorter horizons (e.g., intraday; see Section 2).

So, conceptually, if the underlying quantity (e.g., book value or earnings) has a long time horizon (*i.e.*, changes, say, quarterly), then the corresponding risk factors are not relevant for shorter horizon strategies (e.g., those involving overnight returns),[23] unless there is a linear *extrapolating* argument that reasonably relates such longer term quantities to their shorter term counterparts (as, e.g., in the case of liquidity). More technically, suppose we have K factors we know add value. How do we determine if a new, $(K+1)$-th, factor adds value?[24] Here is a simple method.

Thus, suppose we have N stocks and we have FLM Ω_{iA}, $i = 1, \ldots, N$, $A = 1, \ldots, K$. Let $\Omega'_{iA'}$, $i = 1, \ldots, N$, $A' = 1, \ldots, K'$, $K' \equiv K + 1$ be new FLM once we add a new, $(K+1)$-th, risk factor. (So we have $\Omega'_{iA'}|_{A'=A} = \Omega_{iA}$, $A = 1, \ldots, K$.) Let R_i be the returns used in our strategy, *i.e.*, these returns

---

[19] Similarly, growth does not add value in this context either. This is not to say that, e.g., earnings are not important in short-term trading. However, the way to implement them is via monitoring earnings announcements and, e.g., not trading stocks immediately following their earnings announcements, not by using growth style factor in, say, intraday regressions or optimization.

[20] Arguably, there might be higher-order indirect effects via the book value affecting liquidity and market cap (see below). However, such higher-order effects are expected to be lost in all the noise. They might be ephemerally amplified around the time book value is updated (quarterly).

[21] Market cap is relevant primarily because it is highly correlated with liquidity.

[22] One can directly measure intraday liquidity based on "micro" quantities, which is more tedious. Typically, ADDV based computation reasonably agrees with such "micro" computation.

[23] Conversely, value-based longer horizon strategies would not benefit from any risk factors based on "micro" quantities with, say, millisecond horizons. e.g., statistical arbitrage strategies have high turnover as they attempt to capture intraday mean-reversion effects due to market over-/under-reactions to news events, etc. Value based strategies have very low turnover given that periods of extreme mispricings seldom occur (e.g., '87 Crash, '08 Meltdown).

[24] In the next section we discuss why no-value-adding factors can increase trading costs.

have the time horizon relevant to our strategy.[25] We can run two regressions (without intercept—unless it is included in Ω_{iA}), first R_i over Ω_{iA}, and second R_i over $\Omega'_{iA'}$. In R notations:

$$R \sim -1 + \Omega \tag{54}$$

$$R \sim -1 + \Omega' \tag{55}$$

In actuality, R_i, Ω_{iA} and $\Omega'_{iA'}$ are time series: $R_i(t_s) \equiv R_{si}$, $\Omega_{iA}(t_s) \equiv \Omega_{siA}$, $\Omega'_{iA'}(t_s) \equiv \Omega'_{siA'}$, $s = 0, 1, \ldots, M$. We can run the above two regressions for each value of s and look at, e.g., two time-series vectors of the regression F-statistic to assess if the new risk factor improves the overall F-statistic.[26] Alternatively, we can pull the $(M + 1) \times N$ matrix R_{si} into a vector $\widehat{R}_\sigma$ of length $(M + 1)N$ (*i.e.*, treat the index pair (s, i) as a single index σ), and do the same with FLM: $\Omega_{siA} \equiv \widehat{\Omega}_{\sigma A}$, $\Omega'_{siA'} \equiv \widehat{\Omega}'_{\sigma A'}$. We can now run two regressions

$$\widehat{R} \sim -1 + \widehat{\Omega} \tag{56}$$

$$\widehat{R} \sim -1 + \widehat{\Omega}' \tag{57}$$

and compare the F-statistic.[27] If K is not large, it is also informative to compare the t-values of the regression coefficients and assess the effect of the new factor.

For illustrative purposes, we ran such regressions for overnight returns $R_i \equiv \log(P_i(t_{\text{open}})/P_i(t_{\text{close}}))$, where the open $P_i(t_{\text{open}})$ and the previous close $P_i(t_{\text{close}})$ prices are adjusted for splits and dividends. In the case of, say, book value, as a benchmark it suffices to consider a $K = 1$ model, where the sole risk factor is the intercept. Then we add the second risk factor, which is (log of) book (or tangible book, price-to-book, etc.),[28] so $K' = 2$. The regression F-statistic and t-values are given in Table 1 (for regressions (56) and (57)), which shows that the second regression (57) involving (tangible) book value does not have improved statistic over the intercept-only regression. The 1-factor regressions other than the intercept-only regression can be thought of as regressions over "betas". The $\log(\text{Prc/Book})$ case (see Table 1) is the closest to the intercept-only case because the regression Prc $\sim -1 + $ Book has F-statistic 56,230, and the t-value 237.1, *i.e.*, price and book value are highly correlated. As to the 2-factor regressions, (T)Book does not improve the statistic. It is log(Prc) that makes impact, precisely because prices change daily.

We also ran the (56) and (57) regressions with a $K = 10$ model as a benchmark, where the risk factors are 10 BICS sectors[29] (so $K' = 11$). The results are given in Table 2, which shows that book value does not improve regression statistic. As above, it is log(Prc) that provides improvement. We also ran the (54) and (55) regressions separately for each date (*i.e.*, without pulling the index pair (s, i) into a single index

---

[25] e.g., R_i are overnight returns, we obtain alphas from these returns by regressing them (possibly, with some weights) over some FLM, and then we trade on these alphas right after the open.

[26] To improve statistical significance, outliers can be removed (or smoothed, e.g., Winsorized).

[27] When assessing F-statistic, it needs to be taken into account that we have $(K + 1)$ *vs.* K factors, as is a possible change in the number of observations per factor due to any NAs.

[28] We used fundamental data from stockpup.com (accessed 07/28/2014) and pricing data from finance.yahoo.com (accessed 07/29/2014) from 06/18/2009 through 06/20/2014 for a universe of 493 stocks, essentially from S&P500. Negative (tangible) book values were omitted for the entire backtesting period.

[29] Stocks rarely jump industries let alone sectors, so sector assignments are robust against time.

σ—see above) with the same $K = 10$ benchmark. The results are given in Table 3 and agree with those in Table 2. Log(Prc), not Book, has impact.

Table 1. Results for regressions (56) and (57) with the intercept-only 1-factor model as the benchmark. Int = intercept; (T)Book = (tangible) book; Prc = adjusted previous close; RPrc = raw (unadjusted) previous close. Next to Int+log(Prc) we also give Int+log(RPrc) results. We do this because adjusting the previous close introduces a bias of anticipating future splits and/or dividends. However, as can be seen from the Int+log(RPrc) row, this bias is relatively mild and does not affect our conclusions. The blank entries "—" stand for N/As.

Regression/Statistic	F-Statistic	Intercept *t*-Value	Second Coefficient *t*-Value
Int only	737.7	27.16	—
Book only	237.2	—	15.40
TBook only	191.2	—	13.83
Prc only	1.34	—	1.16
Prc/Book only	12.5	—	3.54
Prc/Tbook only	3.84	—	1.96
log(Book) only	707.5	—	26.60
log(TBook) only	583.7	—	24.70
log(Prc) only	526.0	—	22.94
log(Prc/Book) only	739.1	—	–27.19
log(Prc/Tbook) only	608.7	—	–24.67
Int+Book	362.5	22.08	4.10
Int+TBook	297.6	20.10	4.56
Int+(Prc/Book)	354.3	26.38	–0.66
Int+(Prc/TBook)	287.2	23.89	0.15
Int+Prc	368.9	27.14	–0.24
Int+log(Book)	354.2	0.98	0.53
Int+log(TBook)	294.1	–2.11	3.70
Int+log(Prc)	473.9	20.53	–14.48
Int+log(RPrc)	468.7	20.18	–14.12
Int+log(Prc/Book)	394.0	–6.99	–8.93
Int+log(Prc/TBook)	329.9	–7.14	–9.23

Finally, for the (54) and (55) regressions we computed the t-statistic of actual risk factor time series *a la* Fama and MacBeth [29], both for the $K = 1$ (intercept only) and $K = 10$ (BICS sectors) benchmark factor models. The results are given in Table 4 and Table 5 and agree with those in Table 1, Table 2 and Table 3.

Table 2. Results for regressions (56) and (57) with the BICS-sector 10-factor model as the benchmark. S = 10 BICS sectors labeled by S1(30), S2(63), S3(45), S4(30), S5(91), S6(75), S7(42), S8(48), S9(41) and S10(28) (the parentheticals show the number of tickers in each sector); X = the 11th factor (P, B, P/B, log(P), log(B) and log(P/B)); P = adjusted previous close; B = book; F = F-statistic; t = t-value. e.g., in the "Reg:" line "S+(P/B)" means that the returns R are regressed over FLM Ω containing 11 columns corresponding to the 10 sectors S1 through S10 plus the 11th factor X, which is (P/B) in this case. In the S+log(P) column we also give the values when P is taken to be the raw (unadjusted) previous close. We do this because adjusting the previous close introduces a bias of anticipating future splits and/or dividends. However, as can be seen from the S+log(P) column, this bias is relatively mild and does not affect our conclusions.

Reg:	S	S+P	S+B	S+(P/B)	S+log(P)	S+log(B)	S+log(P/B)
F	80.6	73.3	72.4	71.3	92.6/91.7	71.3	77.8
t:S1	6.40	6.40	6.16	6.40	15.07/14.80	1.56	−6.42
t:S2	6.67	6.67	5.94	6.50	15.74/15.44	1.19	−7.08
t:S3	5.57	5.57	5.02	5.60	15.08/14.76	1.56	−7.04
t:S4	5.40	5.40	5.08	5.41	14.13/13.87	1.32	−6.79
t:S5	13.31	13.28	11.12	13.36	19.26/18.93	1.80	−6.38
t:S6	13.13	13.13	12.26	12.50	19.29/18.99	1.98	−6.09
t:S7	4.97	4.97	3.25	3.74	14.40/14.15	0.89	−7.37
t:S8	6.85	6.85	6.42	6.88	15.88/15.58	1.35	−6.80
t:S9	12.83	12.84	11.90	12.92	19.40/19.12	2.49	−5.39
t:S10	8.63	8.63	8.21	8.87	16.17/15.92	2.19	−5.69
t:X	—	−0.42	3.42	−0.45	−14.57/−14.21	−0.20	−8.47

Table 3. Results for regressions (54) and (55) with the BICS-sector 10-factor model as the benchmark. The notations are the same as in Table 2, except that F = median F-statistic, and t = median t-value, where F-statistic and t-values are computed based on regressions (54) and (55) for each date, and the median is computed serially over all dates. The meaning of double entries in the S+log(P) column is the same as in Table 2.

Reg:	S	S+P	S+B	S+(P/B)	S+log(P)	S+log(B)	S+log(P/B)
F	13.5	12.2	12.0	12.0	12.3/12.3	12.1	12.2
t:S1	0.40	0.40	0.40	0.43	0.67/0.68	0.11	−0.18
t:S2	0.61	0.61	0.62	0.63	0.85/0.83	0.11	−0.16
t:S3	0.34	0.33	0.32	0.32	0.69/0.63	0.11	−0.20
t:S4	0.23	0.23	0.23	0.25	0.62/0.56	0.10	−0.21
t:S5	0.91	0.91	0.76	0.88	0.90/0.85	0.15	−0.19
t:S6	0.80	0.80	0.82	0.91	0.96/0.92	0.14	−0.14
t:S7	0.39	0.39	0.24	0.25	0.70/0.64	0.10	−0.20
t:S8	0.54	0.54	0.57	0.57	0.76/0.74	0.08	−0.19
t:S9	0.76	0.76	0.77	0.80	1.00/0.98	0.18	−0.12
t:S10	0.53	0.53	0.52	0.55	0.84/0.88	0.14	−0.14
t:X	—	−0.02	0.13	−0.04	−0.55/−0.52	−0.03	−0.30

Table 4. Results for regressions (54) and (55) with the intercept-only 1-factor model as the benchmark. The notations are the same as in Table 1, except that the t-statistic here refers to the t-statistic of the corresponding risk factor time series *a la* Fama and MacBeth [29]. These t-statistic are annualized, *i.e.*, we compute the daily t-statistic and then multiply it by $\sqrt{252}$.

Regression/Statistic	Intercept *t*-Statistic	Second Coefficient *t*-Statistic
Int only	0.90	—
Int+Book	0.82	2.21
Int+(Prc/Book)	0.90	–0.69
Int+Prc	0.90	–0.42
Int+log(Book)	0.23	0.32
Int+log(Prc)	1.90	–3.50
Int+log(RPrc)	1.78	–3.10
Int+log(Prc/Book)	–2.15	–2.90

Table 5. Results for regressions (54) and (55) with the BICS-sector 10-factor model as the benchmark. The notations are the same as in Table 2, except that "t:⋆" refers to the annualized t-statistic of the corresponding risk factor "⋆" time-series *a la* Fama and MacBeth [29], same as in Table 4. The meaning of double entries in the S+log(P) column is the same as in Table 2.

Reg:	S	S+P	S+B	S+(P/B)	S+log(P)	S+log(B)	S+log(P/B)
t:S1	0.76	0.76	0.74	0.78	1.76/1.64	0.39	–2.02
t:S2	0.59	0.59	0.54	0.60	1.61/1.50	0.28	–2.24
t:S3	0.76	0.76	0.68	0.77	2.07/1.91	0.30	–2.42
t:S4	1.09	1.09	1.01	1.10	2.32/2.14	0.37	–2.48
t:S5	0.88	0.88	0.77	0.88	1.75/1.65	0.45	–2.02
t:S6	1.07	1.07	1.02	1.05	2.01/1.88	0.51	–1.86
t:S7	0.83	0.83	0.56	0.66	2.14/1.99	0.22	–2.66
t:S8	0.64	0.64	0.60	0.65	1.72/1.59	0.32	–2.07
t:S9	1.09	1.09	1.02	1.09	1.87/1.77	0.61	–1.51
t:S10	1.17	1.17	1.13	1.23	2.04/1.92	0.58	–1.78
t:X	—	–0.56	2.15	–0.61	–3.45/–3.05	0.02	–3.12

4. Pitfalls of Standardized Risk Models

4.1. Industry Risk Factors

Suppose we have an industry classification. For our discussion below it will make no difference whether FLM Ω_{iA}^{ind}, $A = 1, \ldots, K_{\text{ind}}$ is binary or some conglomerates are allowed, so for simplicity we will assume it to be binary (see Subsection 2.1):

$$\Omega_{iA}^{\text{ind}} = \delta_{G(i),A} \tag{58}$$

For definiteness, let us fix the names of the industry tree levels as "sectors → sub-sectors → industries", so "industries" correspond to the most detailed level. The number of the industries K_{ind} depends on the

universe. Different universes U_1 and U_2 can have vastly different industries to which the corresponding stocks belong.

In SRM a large number of stocks (e.g., several thousand for the U.S. models) are "squeezed" into a relatively modest number K_{ind}^* of standardized industries, which can be substantially smaller than the number of true industries K_{ind} for a typical quantitative trading portfolio universe of, say, 1000–2500 names. So, standardized industries typically lose granularity, which determines how well RM helps hedge a portfolio against industry risk. For illustrative purposes, let us look at the number of true industries for top-by-market-cap portfolios in BICS.[30] We require at least 10 stocks in each industry. Small industries are pruned to the sub-sector level and, if need be, to the sector level. Any leftover small industries can be merged into larger industries. The result is given in Table 6. The numbers of true industries for top 1,500+ universes are sizably higher than those of typical standardized industries.

Table 6. Number of BICS industries for portfolios of stocks in top X by market cap with at least 10 stocks in each industry as of August 19, 2014. Only U.S. listed common stocks and class shares are included (no OTCs, preferred shares, etc.).

Top X by Market Cap	Number of Industries
1000	55
1500	75
2000	94
2500	107
3000	122
3500	125
4000	128
4500	130
5000	133

4.2. Empty Standardized Industries

Furthermore, for any given universe U, even if U is, say, 1,000-2,500 names, we can have (almost) empty standardized industries—e.g., a portfolio that does not trade any stocks from a given (sub-)sector. Empty standardized industries would have been omitted had we built RM based on the custom universe U. In SRM we have no such option, so we must keep empty industries. Why is this so bad?

Style factors are not important for our discussion here, so let us consider RM with only industry factors. Let these be K standardized binary industries. Let the risk model universe be U_{SRM}, and let our universe be $U \subset U_{SRM}$. Let FLM be Ω_{iA}, $i \in \{1, \ldots, N\} \equiv U$. Let some industries be empty with $N_A = 0$, where

$$N_A \equiv \sum_{i=1}^{N} \Omega_{iA} = \sum_{i=1}^{N} \delta_{G(i),A} \tag{59}$$

---

[30] BICS naming convention is "sectors $\rightarrow$ industries $\rightarrow$ sub-industries", so our "industries" correspond to BICS "sub-industries", and our "sub-sectors" correspond to BICS "industries".

Such industries must be omitted from regressions, if this is how we use RM.

On the other hand, suppose we are doing optimization, and we cannot omit such empty industries. Let $J \equiv \{A | N_A > 0\}$ and $J' \equiv \{A | N_A = 0\}$. We have

$$\Gamma_{ij} = \xi_i^2 \, \delta_{ij} + \sum_{A,B \in J} \Omega_{iA} \, \Omega_{jB} \, \Phi_{AB} = \xi_i^2 \, \delta_{ij} + \Phi_{G(i),G(j)} \tag{60}$$

The number of risk factors in this model is $\widehat{K} = |J| < K$, yet FCM Φ_{AB} is computed based on K factors. So, from the viewpoint of the universe U, there are $|J'| = K - \widehat{K}$ "hidden" factors, to which U has no exposure, yet they affect the covariance matrix Γ. This does not bode well with the RM premise that all covariances are explained by a combination of: i) some fixed risk factors, exposure to which for a given universe of stocks is well-defined; and ii) ISR, which describes all uncertainty not described by the risk factors. Based on this premise, the correct way of modeling risk for our universe U would be to assume that we have $\widehat{K}$ risk factors and compute FCM $\widehat{\Phi}_{AB}$, $A, B \in J$ for these factors along with the corresponding ISR $\widehat{\xi}_i$, *i.e.*, to have

$$\widehat{\Gamma}_{ij} \equiv \widehat{\xi}_i^2 \, \delta_{ij} + \sum_{A,B \in J} \Omega_{iA} \, \Omega_{jB} \, \widehat{\Phi}_{AB} = \widehat{\xi}_i^2 \, \delta_{ij} + \widehat{\Phi}_{G(i),G(j)} \tag{61}$$

At first it might appear that (60) and (61) are identical—if Φ_{AB} and $\widehat{\Phi}_{AB}$ are computed as SCM of the corresponding risk factors, then we should have $\widehat{\Phi}_{AB} = \Phi_{AB}$, $A, B \in J$. However, as mentioned in Subsection 2.3, in real life FCM is *not* computed as SCM, because ISR with such a computation are typically ill-defined. Consequently, $\widehat{\Phi}_{AB} \neq \Phi_{AB}$, $A, B \in J$ and $\widehat{\xi}_i \neq \xi_i$. Because of this interdependency of FCM and ISR, it is more desirable to compute $\widehat{\Phi}_{AB}$ and $\widehat{\xi}_i$ directly based on the universe U without any empty industries—the latter only bring more noise into the computation. Any uncertainty not described by the relevant risk factors should be modeled via ISR, not some "hidden" risk factors.

That empty industries, to which the universe U has no exposure, add noise can be seen in the optimization context—they contribute to the desired dollar holdings (24) via (25).[31] The effect is that the desired dollar holdings are approximately neutralized against these "hidden" industries (see (26) and (27)).[32] Typically, this generates additional noise ("twitch") trades, which on paper may appear harmless,[33] but in real life can increase trading costs and reduce profitability, rendering empty industries undesirable. The same argument applies to irrelevant style factors (e.g., value/growth in short horizon strategies; see Section 3.2), rendering them harmful.[34]

---

[31] FCM and ISR must be recomputed based on non-empty industries to remove this contribution.

[32] These "hidden" industries might be correlated with the non-empty industries. However, such correlations should not be high—if they are high, then the industry classification is too granular (or deficient) and must be pruned (or replaced with a more precise industry classification). Including redundant noise-generating industries in RM is certainly not the right way to handle such cases.

[33] On paper such noise trades typically have little effect on the simulated P&L, but can reduce the Sharpe ratio—if the approximate neutrality constraints effected by these empty industries do not add value, then the portfolio is suboptimal, *i.e.*, it does not maximize the Sharpe ratio.

[34] This brings us to another point we made in Introduction: in SRM typically FCM is computed based on some universe U_*, a fraction of U_{SRM}. For the same reasons as above, it is preferable to compute FCM based on the trading universe U as U_* may not have substantial overlap with U.

4.3. Style Risk Factors

Some —albeit not necessarily all[35]— style factors are normalized. One way of normalizing a style factor labeled by $A \in \{1, \ldots, K\}$ is by conforming the values of the A-th column in FLM Ω_{iA} to a normal distribution with, e.g., mean 0 and standard deviation equal to the standard deviation (or MAD) of the original, unnormalized column. This is done in the R code in Appendix A for momentum, size and liquidity. If the distribution is expected to be log-normal, then one normalizes log of the column and re-exponentiates. This is done for intraday volatility in the R code in Appendix A.

Such normalizations of style factors are typically done across the entire coverage universe U_{SRM}. Suppose we wish to use SRM for our universe U, which is a fraction of U_{SRM}. Then the values of the corresponding columns in Ω_{iA} with $i \in U$ are no longer normalized. For a random subset $U \subset U_{SRM}$ one may expect that they are still approximately normalized if the number of stocks in U is large. However, typically there is nothing random about a real-life trading universe U, which is carefully selected based on certain requirements on market cap, liquidity, volatility, price and/or other relevant quantities, which can be rather skewed, so the truncated to U style factor columns in Ω_{iA} are no longer even approximately normalized. Why does this matter? If FCM is computed based on normalized style factors (be it using U_{SRM} or U_*), then this FCM is not the same (or even necessarily close to) FCM one would obtain based on style factors normalized based on U. This also affects ISR. Therefore, for the same reasons as in the previous subsection, it is more desirable to compute FLM, FCM and ISR based on the custom universe U.

5. Concluding Remarks

Above we discussed points (1)–(4) raised in Introduction in relation to using SRM and when and why CRM are warranted. Let us conclude by briefly touching upon points (5)–(7) mentioned in Introduction, which relate to other aspects.

Statistical Arbitrage (StatArb) "refers to highly technical short-term mean-reversion strategies involving large numbers of securities (hundreds to thousands, depending on the amount of risk capital), very short holding periods (measured in days to seconds), and substantial computational, trading, and information technology (IT) infrastructure" [39]. A quantitative framework for this "mean-reversion" was recently discussed in [50]. Schematically, one can think about this mean-reversion as follows. Pick some returns. Pick RM. Then: (i) either regress your returns over FLM (or a subset of its columns) with some weights, or (ii) do optimization using RM (possibly with some constraints). One can hang various bells and whistles onto a strategy constructed this way, which do contribute into differentiating models. Nonetheless, the choice of RM is a major factor (along with the choice of returns) in what the desired holdings and trades will look like. As discussed in Section 2.4.3, in optimization desired holdings are approximately neutral w.r.t the risk factors, while in regression they are exactly neutral. So RM, in its uses discussed above, factors away (completely or approximately) the risk exposure in

---

[35] e.g., style factors with discrete values are not normalized. An example is a binary style factor in some SRM indicating whether the stock belongs to the universe U_* defined above.

the returns w.r.t. the risk factors. Therefore, using the same RM for two sets of apparently different returns can substantially reduce the difference between the two—when it comes to the resulting desired holdings—making the two strategies more correlated.

Therefore, to diversify strategies it is not only important to diversify the returns, but also to diversify RM. This is especially useful in the "herd effect" situations, where many market participants—for whatever (temporary) underlying reason (*cf.* Aug'07 Quant Meltdown)—are "compelled" to do the same thing. When it comes to unpleasant situations such as potential book liquidation, being less correlated with others can make a difference between liquidating (and incurring huge transaction costs) or weathering the storm. Custom RM in such cases can make a difference.

We can take this further by noting that, even with the same returns, two substantially different RM, call them RM-A and RM-B, can produce P&L streams which are not that highly correlated. In this case, instead of running one strategy with the returns, say, optimized using only RM-A, it makes more sense to run two strategies—seamlessly, with trades crossed internally at the level of desired holdings, *i.e.*, the two strategies are combined with some weights w_A and w_B (see below)—where the first strategy, call it Str-A, is optimized using RM-A, and the second strategy, call it Str-B, is optimized using RM-B. If both strategies have positive returns and are not too highly correlated, then even if Str-B has worse return than Str-A, it still makes sense to combine them with some weights. In the zeroth approximation the weights w_A and w_B can be obtained, e.g., by requiring that the Sharpe ratio of the resulting combined strategy be maximized. However, in real life—so long as the Sharpe ratio is acceptably high—typically it is the P&L that matters. In this regard, combining the two strategies can increase the P&L as the capacity bound of the combined strategy is higher (than those of Str-A and Str-B)—the Str-A and Str-B trades are not highly correlated and by combining them one reduces turnover, thereby decreasing market impact and increasing capacity, and also decreasing costs of trading. Turnover reduction and increasing capacity are two important incentives for diversifying and using CRM.

Finally, let us mention that CRM provide additional evident benefits listed in point (7) in Introduction. Also, for shorter-term applications (which is precisely when CRM are warranted), which do not require long lookbacks or nontrivial fundamental data (such as earnings going back many years), the data required for building a CRM is typically already available to the portfolio manager. All in all, it boils down to computing FCM and ISR in a self-consistent fashion.

Acknowledgments

A. R Code for Some Style Risk Factors

Below we give R code for 4 style factors (function `calc.add.fac()`): momentum, liquidity, size and intraday volatility. All uninformative dependencies have been omitted. The code is unaltered otherwise. (Functions `normalize()`, `calc.sr()`, `calc.eff.mad()`, `calc.ret.mv()` and `calc.ret.mv.clean()` are auxiliary.) Below: `is.adr` is a binary N-vector (N is the number of stocks); `hist.prc`, `hist.vol`, `hist.high`, `hist.low`, `hist.cap` are $d \times N$ matrices (d is the number of trading dates in the historical data)—historical closing price, daily volume, daily

high, daily low (all adjusted—this is necessary for momentum, albeit not for liquidity or intraday volatility), and market cap, respectively; `dates` is the vector of the last 252 trading dates. ADRs are normalized independently. `days` is the lookback (e.g., 252 trading days), while `back` is for out-of-sample backtesting (and is 0 for current-date calculations); both `days` and `back` are passed into `calc.add.fac()` via the omitted arguments.

```
normalize <- function(x, center = mean(x), sdev = sd(x)){
  return(qnorm(ppoints(x)[sort.list(sort.list(x), method = "radix")],
    center, sdev))
}
calc.sr <- function(tv){
  sr <- sqrt(tv)
  sr <- log(sr)
  sr <- normalize(sr, median(sr), mad(sr))
  sr <- exp(sr)
  return(sr)
}
calc.eff.mad <- function(ret){
  eff.mad <- (5 * outer(apply(ret, 1,
    mad, na.rm = T), apply(ret,
  2, mad, na.rm = T), pmax))
  return(eff.mad)
}
calc.ret.mv <- function(prc, back, days, d.r){
  last <- nrow(prc) - back
  first <- last - days
  today <- prc[last:first, ]
  yest <- 0
  for(i in 1:d.r)
    yest <- yest + prc[(last - i):(first - i), ]
  yest <- yest / d.r
  ret <- today/yest - 1
  dimnames(ret) <- dimnames(prc[last:first, ])
  return(ret)
}
calc.ret.mv.clean <- function(prc, back, days, d.r){
  ret <- calc.ret.mv(prc, back, days, d.r)
  eff.mad <- calc.eff.mad(ret)
  bad <- abs(ret - apply(ret, 1, median)) > eff.mad
  ret[bad] <- NA
  avg.ret <- matrix(rowMeans(ret, na.rm = T),
```

```
    nrow = nrow(ret), ncol = ncol(ret))
  ret[bad] <- avg.ret[bad]
  ret <- ret - avg.ret
  return(ret)
}
calc.add.fac <- function(...){
  ### MOMENTUM MOVING AVERAGE LENGTHS    d.r <- 5

  ### ADDV MOVING AVERAGE LENGTHS
  d.addv <- 20

  ### MOMENTUM FACTOR
  ### BASED ON AVERAGE 5-DAY RETURNS (OUTLIERS REMOVED)
  ret.mom <- calc.ret.mv.clean(hist.prc, back, days, d.r)
  mom <- apply(ret.mom, 2, mean)
  mom <- normalize(mom, 0, mad(mom))

  ### AVERAGE DAILY DOLLAR VOLUME (ADDV) FACTOR
  ### BASED ON LAST 20 DAYS
  ### ADRS ARE NORMALIZED ACCORDING TO NON-ADR DISTRIBUTION
  not.adr <- !is.adr
  addv <- hist.prc[dates, ] * hist.vol[dates, ]
  addv <- addv[1:d.addv, ]
  addv[addv == 0] <- NA
  addv <- colMeans(log(addv), na.rm = T)
  addv[is.adr] <- normalize(addv[is.adr], 0, mad(addv[not.adr]))
  addv <- normalize(addv, 0, mad(addv[not.adr]))

  ### MARKET CAP FACTOR
  ### BASED ON 252 DAYS
  ### ADRS ARE NORMALIZED ACCORDING TO NON-ADR DISTRIBUTION
  mkt.cap <- hist.cap[dates, ]
  mkt.cap[mkt.cap == 0] <- NA
  mkt.cap <- colMeans(log(mkt.cap), na.rm = T)
  mkt.cap[is.adr] <- normalize(mkt.cap[is.adr], 0, mad(mkt.cap[not.adr]))
  mkt.cap <- normalize(mkt.cap, 0, mad(mkt.cap[not.adr]))

  ### INTRADAY VOLATILITY FACTOR
  ### BASED ON 252 DAYS
  hist.low <- hist.low[dates, ]
  hist.high <- hist.high[dates, ]
```

```
  hist.prc <- hist.prc[dates, ]
  high.low <- abs(hist.high - hist.low) / hist.prc
  high.low <- calc.sr(colMeans(high.low^2))
}
```

B. C Code for Symmetric Matrix Inversion

The vector `a[]` is a symmetric $n \times n$ matrix A_{ij} to be inverted pulled into a vector such that A_{ij}, $i, j = 0, 1, \ldots, n - 1$ is given by `a[i + n * j]`. This algorithm utilizes the fact that the matrix is symmetric thereby reducing the number of operations (compared with the Gauss-Jordan method for a general matrix).

```c
static void InvSymMat(double *a, int n){
  int i, j, k;
  double sum;

  for( i = 0; i < n; i++ )
    for( j = i; j < n; j++ )
    {
      sum = a[i + n * j];

      for( k = i - 1; k >= 0; k- )
        sum -= a[i + n * k] * a[j + n * k];

      a[j + n * i] =
        ( j == i ) ?  1 / sqrt(sum) :  sum * a[i * (n + 1)];
    }

  for( i = 0; i < n; i++ )
    for( j = i + 1; j < n; j++ )
    {
      sum = 0;

      for( k = i; k < j; k++ )
        sum -= a[j + n * k] * a[k + n * i];

      a[j + i * n] = sum * a[j * (n + 1)];
    }

  for( i = 0; i < n; i++ )
    for( j = i; j < n; j++ )
    {
```

```
      sum = 0;
      for( k = j; k < n; k++ )
        sum += a[k + n * i] * a[k + n * j];


      a[i + n * j] = a[j + n * i] = sum;
    }
}
```

C. Disclaimers

Wherever the context so requires, the masculine gender includes the feminine and/or neuter, and the singular form includes the plural and *vice versa*. The author of this paper ("Author") and his affiliates including without limitation Quantigic® Solutions LLC ("Author's Affiliates" or "his Affiliates") make no implied or express warranties or any other representations whatsoever, including without limitation implied warranties of merchantability and fitness for a particular purpose, in connection with or with regard to the content of this paper including without limitation any code or algorithms contained herein ("Content").

The reader may use the Content solely at his/her/its own risk and the reader shall have no claims whatsoever against the Author or his Affiliates and the Author and his Affiliates shall have no liability whatsoever to the reader or any third party whatsoever for any loss, expense, opportunity cost, damages or any other adverse effects whatsoever relating to or arising from the use of the Content by the reader including without any limitation whatsoever: any direct, indirect, incidental, special, consequential or any other damages incurred by the reader, however caused and under any theory of liability; any loss of profit (whether incurred directly or indirectly), any loss of goodwill or reputation, any loss of data suffered, cost of procurement of substitute goods or services, or any other tangible or intangible loss; any reliance placed by the reader on the completeness, accuracy or existence of the Content or any other effect of using the Content; and any and all other adversities or negative effects the reader might encounter in using the Content irrespective of whether the Author or his Affiliates is or are or should have been aware of such adversities or negative effects.

The R code included in Appendix A hereof is part of the copyrighted R code for Quantigic® Risk Model™ and is provided herein with the express permission of Quantigic® Solutions LLC. The copyright owner retains all rights, title and interest in and to its copyrighted source code included in Appendix A and Appendix B hereof and any and all copyrights therefor.

The Content is not intended, and should not be construed, as an investment, legal, tax or any other such advice, and in no way represents views of Quantigic® Solutions LLC, the website: www.quantigic.com or any of their other affiliates.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Banz, R. The relationship between return and market value of common stocks. *J. Financ. Econ.* **1981**, *9*, 3–18.
2. Basu, S. The investment performance of common stocks in relation to their price to earnings ratios: A test of the efficient market hypothesis. *J. Financ.* **1977**, *32*, 663–682.
3. Fama, E.; French, K. The cross-section of expected stock returns. *J. Financ.* **1992**, *47*, 427–465.
4. Fama, E.F.; French, K.R. Common risk factors in the returns on stocks and bonds. *J. Financ. Econ.* **1993**, *33*, 3–56.
5. Lakonishok, J.; Shleifer, A.; Vishny, R.W. Contrarian investment, extrapolation, and risk. *J. Financ.* **1994**, *49*, 1541–1578.
6. Liew, J.; Vassalou, M. Can book-to-market, size and momentum be risk factors that predict economic growth? *J. Financ. Econ.* **2000**, *57*, 221–245.
7. Asness, C.; Stevens, R. *Intra- and Inter-Industry Variation in the Cross-Section of Expected Stock Returns. Working Paper*; Goldman Sachs, Investment Management Division: San Fransisco, CA, USA, 1995.
8. Haugen, R.A. *The New Finance: The Case Against Efficient Markets*; Prentice Hall: Upper Saddle River, NJ, USA, 1995.
9. Asness, C.S. *The Power of Past Stock Returns to Explain Future Stock Returns. Working Paper*; Goldman Sachs, Investment Management Division: San Fransisco, CA, USA, 1995.
10. Jegadeesh, N.; Titman, S. Returns to buying winners and selling losers: Implications for stock market efficiency. *J. Financ.* **1993**, *48*, 65–91.
11. Asness, C.; Krail, R.J.; Liew, J.M. Do hedge funds hedge? *J. Portf. Manag.* **2001**, *28*, 6–19.
12. Anson, M. Performance measurement in private equity: Another look at the lagged beta effect. *J. Priv. Equity* **2013**, *17*, 29–44.
13. Pastor, L.; Stambaugh, R.F. Liquidity risk and expected stock returns. *J. Polit. Econ.* **2003**, *111*, 642–685.
14. Scholes, M.; Williams, J. Estimating betas from nonsynchronous data. *J. Financ. Econ.* **1977**, *5*, 309–327.
15. Ang, A.; Hodrick, R.; Xing, Y.; Zhang, X. The cross-section of volatility and expected returns. *J. Financ.* **2006**, *61*, 259–299.
16. Jagannathan, R.; Wang, Z. The conditional CAPM and the cross-section of expected returns. *J. Financ.* **1996**, *51*, 3–53.
17. Black, F.; Jensen, M.; Scholes, M. The capital asset pricing model: Some empirical tests. In *Studies in the Theory of Capital Markets*; Jensen, M., Ed.; Praeger Publishers: New York, NY, USA, 1972; pp. 79–121.
18. Black, F. Capital market equilibrium with restricted borrowing. *J. Bus.* **1972**, *45*, 444–455.
19. Blume, O.; Friend, L. A new look at the capital asset pricing model. *J. Financ.* **1973**, *28*, 19–33.
20. Brandt, M.W.; Brav, A.; Graham, J.R.; Kumar, A. The idiosyncratic volatility puzzle: Time trend or speculative episodes? *Rev. Financ. Stud.* **2010**, *23*, 863–899.
21. Campbell, J. Stock returns and the term structure. *J. Financ. Econ.* **1987**, *18*, 373–399.

22. Campbell, J.; Shiller, R. The dividend-price ratio and expectations of future dividends and discount factors. *Rev. Financ. Stud.* **1988**, *1*, 195–227.

23. Campbell, J.Y.; Lettau, M.; Malkiel, B.G.; Xu, Y. Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. *J. Financ.* **2001**, *56*, 1–43.

24. Carhart, M.M. Persistence in mutual fund performance. *J. Financ.* **1997**, *52*, 57–82.

25. Chen, N.; Roll, R.; Ross, S. Economic forces and the stock market. *J. Bus.* **1986**, *59*, 383–403.

26. Cochrane, J.H. *Asset Pricing*; Princeton University Press: Princeton, NJ, USA, 2001.

27. Connor, G.; Korajczyk, R. Risk and return in an equilibrium APT: Application of a new test methodology. *J. Financ. Econ.* **1988**, *21*, 255–289.

28. DeBondt, W.; Thaler, R. Does the stock market overreact? *J. Financ.* **1985**, *40*, 739–805.

29. Fama, E.F.; MacBeth, J.D. Risk, return and equilibrium: Empirical tests. *J. Polit. Econ.* **1973**, *81*, 607–636.

30. Fama, E.; French, K. Multifactor explanations for asset pricing anomalies. *J. Financ.* **1996**, *51*, 55–94.

31. Ferson, W.; Harvey, C. The variation in economic risk premiums. *J. Polit. Econ.* **1991**, *99*, 385–415.

32. Ferson, W.; Harvey, C. Conditioning variables and the cross section of stock returns. *J. Financ.* **1999**, *54*, 1325–1360.

33. Hall, A.D.; Hwang, S.; Satchell, S.E. Using bayesian variable selection methods to choose style factors in global stock return models. *J. Bank. Financ.* **2002**, *26*, 2301–2325.

34. Kothari, S.; Shanken, J. Book-to-market, dividend yield and expected market returns: A time series analysis. *J. Financ. Econ.* **1997**, *44*, 169–203.

35. Lee, J.-H.; Stefek, D. Do risk factors eat alphas? *J. Portfolio Manag.* **2008**, *34*, 12–24.

36. Lehmann, B.; Modest, D. The empirical foundations of the arbitrage pricing theory. *J. Financ. Econ.* **1988**, *21*, 213–254.

37. Lintner, J. The valuation of risky assets and the selection of risky investments in stock portfolios and capital budgets. *Rev. Econ. Stat.* **1965**, *47*, 13–37.

38. Lo, A.W.; MacKinlay, A.C. Data-snooping biases in tests of financial asset pricing models. *Rev. Financ. Stud.* **1990**, *3*, 431–468.

39. Lo, A.W. *Hedge Funds: An Analytic Perspective*; Princeton University Press: Princeton, NJ, USA, 2010; p. 260.

40. MacKinlay, A.C. Multifactor models do not explain deviations from the CAPM. *J. Financ. Econ.* **1995**, *38*, 3–28.

41. Merton, R. An intertemporal capital asset pricing model. *Econometrica* **1973**, *41*, 867–887.

42. Mukherjee, D.; Mishra, A.K. Multifactor Capital Asset Pricing Model Under Alternative Distributional Specification. Available online: http://ssrn.com/abstract=871398 (accessed on 7 May 2015).

43. Ng, V.; Engle, R.F.; Rothschild, M. A multi-dynamic-factor model for stock returns. *J. Econ.* **1992**, *52*, 245–266.

44. Ross, S. The arbitrage theory of capital asset pricing. *J. Econ. Theory* **1976**, *13*, 341–360.

45. Schwert, G. Stock returns and real activity: A century of evidence. *J. Financ.* **1990**, *45*, 1237–1257.

46. Sharpe, W. Capital asset prices: A theory of market equilibrium under conditions of risk. *J. Financ.* **1964**, *19*, 425–442.

47. Whitelaw, R. Time variations and covariations in the expectation and volatility of stock market returns. *J. Financ.* **1997**, *49*, 515–541.

48. Zhang, C. A Re-examination of the causes of time-varying stock return volatilities. *J. Financ. Quant. Anal.* **2010**, *45*, 663–684.

49. Harvey, C.R.; Liu, Y.; Zhou, H. ...and the cross-section of expected returns. *Rev. Financ. Stud.* Available online: http://ssrn.com/abstract=2249314 (accessed on 7 May 2015).

50. Kakushadze, Z. Mean-reversion and optimization. *J. Asset Manag.* **2015**, *16*, 14–40.