

Heterotic Risk Models

Zura Kakushadze^{§†1}

[§] *Quantigic[®] Solutions LLC*

*1127 High Ridge Road #135, Stamford, CT 06905*²

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

(April 30, 2015)

Abstract

We give a complete algorithm and source code for constructing what we refer to as heterotic risk models (for equities), which combine: i) granularity of an industry classification; ii) diagonality of the principal component factor covariance matrix for any sub-cluster of stocks; and iii) dramatic reduction of the factor covariance matrix size in the Russian-doll risk model construction. This appears to prove a powerful approach for constructing out-of-sample stable short-lookback risk models. Thus, for intraday mean-reversion alphas based on overnight returns, Sharpe ratio optimization using our heterotic risk models sizably improves the performance characteristics compared to weighted regressions based on principal components or industry classification. We also give source code for: a) building statistical risk models; and ii) Sharpe ratio optimization with homogeneous linear constraints and position bounds.

¹ Zura Kakushadze, Ph.D., is the President of Quantigic[®] Solutions LLC, and a Full Professor at Free University of Tbilisi. Email: zura@quantigic.com

² DISCLAIMER: This address is used by the corresponding author for no purpose other than to indicate his professional affiliation as is customary in publications. In particular, the contents of this paper are not intended as an investment, legal, tax or any other such advice, and in no way represent views of Quantigic[®] Solutions LLC, the website www.quantigic.com or any of their other affiliates. The title of this paper is inspired by the heterotic string theory, with no substantive connection therewith.

1 Introduction

When the number of stocks in a portfolio is large and the number of available (relevant) observations in the historical time series of returns is limited – which is essentially a given for short-horizon quantitative trading strategies – the sample covariance matrix (SCM) is badly singular. This makes portfolio optimization – *e.g.*, Sharpe ratio maximization – challenging as it requires the covariance matrix to be invertible. A standard method for circumventing this difficulty is to employ factor models,³ which, instead of computing SCM for a large number of stocks, allow to compute a factor covariance matrix (FCM) for many fewer risk factors. However, the number of relevant risk factors itself can be rather large. *E.g.*, in a (desirably) granular industry classification (IC), the number of industries⁴ can be in 3 digits for a typical (liquid) trading universe. Then, even the sample FCM can be singular.

In (Kakushadze, 2014b) a simple idea was set forth: model FCM itself via a factor model, and repeat this process until the remaining FCM is small enough and can be computed. In fact, at the end of this process we may even end up with a single factor, for which “FCM” is simply its variance.⁵ This construction – termed in (Kakushadze, 2014b) as “Russian-doll” risk models – dramatically reduces the number of or altogether eliminates the factors for which (off-diagonal) FCM must be computed. The “catch” is that at each successive step we must: i) identify the risk factors; and ii) compute the specific (idiosyncratic) risk (ISR) and FCM consistently.

³ For a partial list of literature related to factor risk models, see, *e.g.*, (Acharya and Pedersen, 2005), (Ang *et al*, 2006), (Anson, 2013/14), (Asness, 1995), (Asness and Stevens, 1995), (Asness *et al*, 2001), (Bai, 2003), (Bai and Li, 2012), (Bai and Ng, 2002), (Bansal and Viswanathan, 1993), (Banz, 1981), (Basu, 1977), (Black, 1972), (Black *et al*, 1972), (Blume and Friend, 1973), (Brandt *et al*, 2010), (Briner and Connor, 2008), (Burmeister and Wall, 1986), (Campbell, 1987), (Campbell *et al*, 2001), (Campbell and Shiller, 1988), (Carhart, 1997), (Chamberlain and Rothschild, 1983), (Chan *et al*, 1985), (Chen *et al*, 1986, 1990), (Chicheportiche and Bouchaud, 2014), (Cochrane, 2001), (Connor, 1984, 1995), (Connor and Korajczyk, 1988, 1989, 2010), (Daniel and Titman, 1997), (DeBondt and Thaler, 1985), (Dhrymes *et al*, 1984), (Fama and French, 1992, 1993, 1996, 2015), (Fama and McBeth, 1973), (Ferson and Harvey, 1991, 1999), (Forni *et al*, 2000, 2005), (Forni and Lippi 2001), (Goyal *et al*, 2008), (Goyal and Santa-Clara, 2003), (Grinold and Kahn, 2000), (Hall *et al*, 2002), (Haugen, 1995), (Heaton and Lucas, 1999), (Heston and Rouwenhorst, 1994), (Jagannathan and Wang, 1996), (Jegadeesh and Titman, 1993, 2001), (Kakushadze, 2014a, 2014b, 2015a), (Kakushadze and Liew, 2014), (King, 1966), (Korajczyk and Sadka, 2008), (Kothari and Shanken, 1997), (Lakonishok *et al*, 1994), (Lee and Stefek, 2008), (Lehmann and Modest, 1988), (Liew and Vassalou, 2000), (Lintner, 1965), (Lo, 2010), (Lo and MacKinlay, 1990), (MacKinlay, 1995), (MacQueen, 2003), (Markowitz, 1952, 1984), (Menchero and Mitra, 2008), (Merton, 1973), (Miller, 2006), (Motta *et al*, 2011), (Mukherjee and Mishra, 2005), (Ng *et al*, 1992), (Pastor and Staambaugh, 2003), (Roll and Ross, 1980), (Rosenberg, 1974), (Ross, 1976, 1978a, 1978b), (Scholes and Williams, 1977), (Schwert, 1990), (Shanken, 1987, 1990), (Shanken and Weinstein, 2006), (Sharpe, 1963, 1964), (Stock and Watson, 2002a, 2002b), (Stroyny, 2005), (Treynor, 1999), (Vassalou, 2003), (Whitelaw, 1997), (Zangari, 2003), (Zhang, 2010), and references therein.

⁴ By this we mean the stock clusters at the most granular level in the IC hierarchy. *E.g.*, in BICS these would be sub-industries, whereas other ICs have different naming conventions.

⁵ Generally, off-diagonal elements of a sample (stock or factor) covariance matrix tend to be unstable out-of-sample, whereas its diagonal elements (variances) typically are much more stable.

Identifying the risk factors in the Russian-doll construction is facilitated by using a binary industry classification:⁶ using BICS as an illustrative example, industries serve as the risk factors for sub-industries; sectors – there are only 10 of them – serve as the risk factors for industries; and – if need be – the “market” serves as the sole risk factor for sectors. Correctly computing ISR and FCM is more nontrivial: the algorithms for this are generally deemed proprietary. One method in the “lore” is to use a cross-sectional linear regression, where the returns are regressed over a factor loadings matrix (FLM), and FCM is identified with the (serial) covariance matrix of the regression coefficients, whereas ISR squared is identified with the (serial) variance of the regression residuals. However, as discussed in (Kakushadze, 2014b), generally this does not satisfy a nontrivial requirement (which is often overlooked in practice) that the factor model reproduce the historical in-sample total variance.

In this paper we share a complete algorithm and source code for building what we refer to as “heterotic” risk models. It is based on a simple observation that, if we use principal components (PCs) as FLM, the aforementioned total variance condition is automatically satisfied. Unfortunately, the number of useful PCs is few as it is limited by the number of observations, and they also tend to be unstable out-of-sample (as they are based on off-diagonal covariances), with the first PC being most stable. We circumvent this by building FLM from the first PCs of the blocks (sub-matrices) of the sample correlation matrix⁷ corresponding to – in the BICS language – the sub-industries. *I.e.*, if there are N stocks and K sub-industries, FLM is $N \times K$, and in each column the elements corresponding to the tickers in that sub-industry are proportional to the first PC of the corresponding block, with all other elements vanishing.⁸ The total variance condition is automatically satisfied. Then, applying the Russian-doll construction yields a nonsingular factor model covariance matrix, which, considering it sizably adds value in Sharpe ratio optimization for certain intraday mean-reversion alphas we backtest, appears to be stable out-of-sample.

Heterotic risk models are based on our proprietary know-how. We hope sharing it with the investment community encourages organic custom risk models building.

This paper is organized as follows. In Section 2 we briefly review some generalities of factor models and discuss in detail the total variance condition. In Section 3 we discuss the PC approach and an algorithm for fixing the number of PC factors, with the R source code in Appendix A. We discuss heterotic risk models in detail in Section 4, with the complete Russian-doll embedding in Section 5 and the R source code in Appendix B. In Section 6 we run a horse race of intraday mean-reversion alphas via i) weighted regressions and ii) optimization using heterotic risk models. For optimization with homogeneous linear constraints and (liquidity/position) bounds we use the R source code in Appendix C.⁹ We briefly conclude in Section 7.

⁶ The number of non-binary style factors is at most of order 10 and does not pose a difficulty for computing the factor covariance matrix. It is the ubiquitous industry factors that are problematic.

⁷ And not SCM – this is an important technical detail, see the discussion in Section 2.4.

⁸ Note that this is not the same as a “hybrid” (mixture) of industry and statistical risk factors.

⁹ The source code given in the appendices is not written to be “fancy” or optimized for speed

2 Multi-factor Risk Models

2.1 Generalities

In a multi-factor risk model, a sample covariance matrix C_{ij} for N stocks, $i, j = 1, \dots, N$, which is computed based on a time series of stock returns R_i (*e.g.*, daily close-to-close returns), is modeled via a constructed covariance matrix Γ_{ij} given by

$$\Gamma \equiv \Xi + \Omega \Phi \Omega^T \quad (1)$$

$$\Xi_{ij} \equiv \xi_i^2 \delta_{ij} \quad (2)$$

where δ_{ij} is the Kronecker delta; Γ_{ij} is an $N \times N$ matrix; ξ_i is the specific risk (a.k.a. idiosyncratic risk) for each stock; Ω_{iA} is an $N \times K$ factor loadings matrix; and Φ_{AB} is a $K \times K$ factor covariance matrix, $A, B = 1, \dots, K$, where $K \ll N$. *I.e.*, the random processes Υ_i corresponding to N stock returns are modeled via N random processes χ_i (specific risk) together with K random processes f_A (factor risk):

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

$$\text{Cov}(\chi_i, \chi_j) = \Xi_{ij} \quad (4)$$

$$\text{Cov}(\chi_i, f_A) = 0 \quad (5)$$

$$\text{Cov}(f_A, f_B) = \Phi_{AB} \quad (6)$$

$$\text{Cov}(\Upsilon_i, \Upsilon_j) = \Gamma_{ij} \quad (7)$$

When $M < N$, where $M + 1$ is the number of observations in each time series, the sample covariance matrix C_{ij} is singular with M nonzero eigenvalues. In contrast, assuming all $\xi_i > 0$ and Φ_{AB} is positive-definite, then Γ_{ij} is automatically positive-definite (and invertible). Furthermore, the off-diagonal elements of C_{ij} typically are not expected to be too stable out-of-sample. On the contrary, the factor model covariance matrix Γ_{ij} is expected to be much more stable as the number of risk factors, for which the factor covariance matrix Φ_{AB} needs to be computed, is $K \ll N$.

2.2 Conditions on Total Variances

The prime aim of a risk model is to predict the covariance matrix out-of-sample as precisely as possible, including the out-of-sample total variances. However, albeit this requirement is often overlooked in practical applications, a well-built factor model had better reproduce the in-sample total variances. That is, we require that the factor model total variance Γ_{ii} coincide with the in-sample total variance C_{ii} :

$$\Gamma_{ii} = \xi_i^2 + \sum_{A,B=1}^K \Omega_{iA} \Phi_{AB} \Omega_{iB} = C_{ii} \quad (8)$$

or in any other way. Its sole purpose is to illustrate the algorithms described in the main text in a simple-to-understand fashion. Some legalese relating to this code is given in Appendix D.

A priori this gives N conditions¹⁰ for $N + K(K + 1)/2$ unknowns ξ_i and Φ_{AB} , so we need additional assumptions¹¹ to compute ξ_i and Φ_{AB} .

2.3 Linear Regression

One such assumption – *intuitively* – is that the total risk should be attributed to the factor risk to the greatest extent possible, *i.e.*, the part of the total risk attributed to the specific risk should be minimized. One way to formulate this requirement mathematically is via least squares. First, mimicking (3), we decompose the stock returns R_i via a linear model

$$R_i = \epsilon_i + \sum_{A=1}^K \Omega_{iA} f_A \quad (9)$$

Here the residuals ϵ_i are *not* the same as χ_i in (3); in particular, generally the covariance matrix $\text{Cov}(\epsilon_i, \epsilon_j)$ is not diagonal (see below). We can require that

$$\sum_{i=1}^N z_i \epsilon_i^2 \rightarrow \min \quad (10)$$

where $z_i > 0$, and the minimization is w.r.t. f_A . This produces a weighted linear regression¹² with the regression weights z_i . So, what should these weights be?

2.4 Correlations, Not Covariances

While choosing unit weights $z_i \equiv 1$ might appear as the simplest thing to do, this suffers from a shortcoming. Intuitively it is clear that – *on average* – the residuals ϵ_i are larger for more volatile stocks, so the regression with unit weights would produce skewed results.¹³ This can be readily rectified using nontrivial regression weights. A “natural” choice is $z_i = 1/C_{ii}$. In fact, we have a regression with unit weights:

$$\tilde{R}_i = \tilde{\epsilon}_i + \sum_{A=1}^K \tilde{\Omega}_{iA} f_A \quad (11)$$

$$\sum_{i=1}^N \tilde{\epsilon}_i^2 \rightarrow \min \quad (12)$$

where $\tilde{R}_i \equiv R_i/\sqrt{C_{ii}}$, $\tilde{\Omega}_{iA} \equiv \Omega_{iA}/\sqrt{C_{ii}}$, and $\tilde{\epsilon}_i \equiv \epsilon_i/\sqrt{C_{ii}}$ on average are expected to be much more evenly distributed compared with ϵ_i – we have scaled away the volatility skewness via rescaling the returns, factor loadings and residuals by $\sqrt{C_{ii}}$.

¹⁰ With additional assumptions not all of these conditions are nontrivial (see below).

¹¹ There are no “natural” $K(K + 1)/2$ conditions we can impose on Γ_{ij} , $i \neq j$ in terms of out-of-sample unstable C_{ij} , $i \neq j$. Note that the variances C_{ii} typically are much more stable.

¹² Without the intercept, that is, unless the intercept is already subsumed in Ω_{iA} .

¹³ Cross-sectionally, stock volatility typically has a roughly log-normal distribution.

So, we are now modeling the sample *correlation* matrix $\Psi_{ij} \equiv C_{ij}/\sqrt{C_{ii}}\sqrt{C_{jj}}$ (note that $\Psi_{ii} = 1$, while $|\Psi_{ij}| \leq 1$ for $i \neq j$)¹⁴ via another factor model matrix

$$\tilde{\Gamma}_{ij} = \tilde{\xi}_i^2 \delta_{ij} + \sum_{A,B=1}^K \tilde{\Omega}_{iA} \Phi_{AB} \tilde{\Omega}_{jB} \quad (13)$$

where $\tilde{\Gamma}_{ij} \equiv \Gamma_{ij}/\sqrt{C_{ii}}\sqrt{C_{jj}}$, and $\tilde{\xi}_i^2 \equiv \xi_i^2/C_{ii}$. The solution to (12) is given by (in matrix notation)

$$f = \left(\tilde{\Omega}^T \tilde{\Omega} \right)^{-1} \tilde{\Omega}^T \tilde{R} \quad (14)$$

$$\tilde{\epsilon} = [1 - Q] \tilde{R} \quad (15)$$

$$Q \equiv \tilde{\Omega} \left(\tilde{\Omega}^T \tilde{\Omega} \right)^{-1} \tilde{\Omega}^T \quad (16)$$

where Q is a projection operator: $Q^2 = Q$. Consequently, we have:

$$\hat{\Xi} \equiv \text{Cov}(\tilde{\epsilon}, \tilde{\epsilon}^T) = [1 - Q] \Psi [1 - Q^T] \quad (17)$$

$$\tilde{\Omega} \Phi \tilde{\Omega}^T = \tilde{\Omega} \text{Cov}(f, f^T) \tilde{\Omega}^T = Q \Psi Q^T \quad (18)$$

Note that the matrix $\hat{\Xi}$ is not diagonal. However, the idea here is to identify $\tilde{\xi}_i^2$ with the diagonal part of $\hat{\Xi}$:

$$\tilde{\xi}_i^2 \equiv \hat{\Xi}_{ii} = ([1 - Q] \Psi [1 - Q^T])_{ii} \quad (19)$$

and we have

$$\tilde{\Gamma}_{ij} = \tilde{\xi}_i^2 \delta_{ij} + (Q \Psi Q^T)_{ij} \quad (20)$$

Note that $\tilde{\xi}_i^2$ defined via (19) are automatically positive (nonnegative, to be precise – see below). However, we must satisfy the conditions (8), which reduce to

$$\tilde{\Gamma}_{ii} = \tilde{\xi}_i^2 + \sum_{A,B=1}^K \tilde{\Omega}_{iA} \Phi_{AB} \tilde{\Omega}_{iB} = \Psi_{ii} = 1 \quad (21)$$

and imply

$$T_{ii} = 0 \quad (22)$$

$$T \equiv 2 Q \Psi Q^T - Q \Psi - \Psi Q^T \quad (23)$$

The N conditions (22) are not all independent. Thus, we have $\text{Tr}(T) = 0$.

¹⁴ If $R_i(t_s)$ (t_s labels the observations in the time series and in the above notations the index s takes $M + 1$ values) are the time series of the stock returns based on which the sample covariance matrix C_{ij} is computed (so $C_{ii} = \text{Var}(R_i(t_s))$, where the variance is serial), then Ψ_{ij} is the sample covariance matrix for the “normalized” returns $\tilde{R}_i(t_s) \equiv R_i(t_s)/\sqrt{C_{ii}}$, i.e., $\Psi_{ij} = \text{Cov}(\tilde{R}_i(t_s), \tilde{R}_j(t_s)) = \text{Cor}(R_i(t_s), R_j(t_s))$, where $\text{Cov}(\cdot, \cdot)$ and $\text{Cor}(\cdot, \cdot)$ are serial.

3 Principal Components

The conditions (22) are nontrivial. They are not satisfied for an arbitrary factor loadings matrix $\tilde{\Omega}_{iA}$. However, there is a simple way of satisfying these conditions, to wit, by building $\tilde{\Omega}_{iA}$ from the principal components of the correlation matrix Ψ_{ij} .

Let $V_i^{(a)}$, $a = 1, \dots, N$ be the N principal components of Ψ_{ij} forming an orthonormal basis

$$\sum_{j=1}^N \Psi_{ij} V_j^{(a)} = \lambda^{(a)} V_i^{(a)} \quad (24)$$

$$\sum_{i=1}^N V_i^{(a)} V_i^{(b)} = \delta_{ab} \quad (25)$$

such that the eigenvalues $\lambda^{(a)}$ are ordered decreasingly: $\lambda^{(1)} > \lambda^{(2)} > \dots$. More precisely, some eigenvalues may be degenerate. For simplicity – and this is not critical here – we will assume that all positive eigenvalues are non-degenerate. However, we can have multiple null eigenvalues. Typically, the number of nonvanishing eigenvalues¹⁵ is M , where, as above, $M+1$ is the number of observations in the stock return time series. We can readily construct a factor model with $K \leq M$:

$$\tilde{\Omega}_{iA} = \sqrt{\lambda^{(A)}} V_i^{(A)} \quad (26)$$

Then the factor covariance matrix $\Phi_{AB} = \delta_{AB}$ and we have

$$\tilde{\Gamma}_{ij} = \tilde{\xi}_i^2 \delta_{ij} + \sum_{A=1}^K \lambda^{(A)} V_i^{(A)} V_j^{(A)} \quad (27)$$

$$\tilde{\xi}_i^2 = 1 - \sum_{A=1}^K \lambda^{(A)} \left(V_i^{(A)} \right)^2 \quad (28)$$

so $\tilde{\Gamma}_{ii} = \Psi_{ii} = 1$. See Appendix B for the R code including the following algorithm.

3.1 Fixing K

When $K = M$ we have $\tilde{\Gamma} = \Psi$, which is singular.¹⁶ Therefore, we must have $K \leq K_{max} < M$. So, how do we determine K_{max} ? And is there K_{min} other than the evident answer $K_{min} = 1$? Here we can do a lot of complicated, even convoluted things. Or we can take a pragmatic approach and come up with a simple heuristic. Here is one simple algorithm that does a very decent job at fixing K .

¹⁵ This number can be smaller if some stock returns are 100% correlated or anti-correlated. For the sake of simplicity – and this not critical here – we will assume that there are no such returns.

¹⁶ Note that $\tilde{\xi}_i^2 = \sum_{a=K+1}^M \lambda^{(a)} \left(V_i^{(a)} \right)^2 \geq 0$. We are assuming $\lambda^{(a)} \geq 0$, which (up to computational precision) is the case if there are no N/As in the stock return times series.

The idea here is simple. It is based on the observation that, as K approaches M , $\min(\tilde{\xi}_i^2)$ goes to 0 (*i.e.*, less and less of the total risk is attributed to the specific risk, and more and more of it is attributed to the factor risk), while as K approaches 0, $\max(\tilde{\xi}_i^2)$ goes to 1 (*i.e.*, less and less of the total risk is attributed to the factor risk, and more and more of it is attributed to the specific risk). So, as a rough cut, we can think of K_{max} and K_{min} as the maximum/minimum values of K such that $\min(\tilde{\xi}_i^2) \geq \zeta_{min}^2$ and $\max(\tilde{\xi}_i^2) \leq \zeta_{max}^2$, where ζ_{min} and ζ_{max} are some desired bounds on the fraction of the contribution of the specific risk into the total risk. *E.g.*, we can set $\zeta_{min} = 10\%$ and $\zeta_{max} = 90\%$. In practice, we actually need to fix the value of K , not K_{max} and K_{min} , especially that for some preset values of ζ_{min} and ζ_{max} we may end up with $K_{max} < K_{min}$. However, the above discussion aids us in coming up with a simple heuristic definition for what K should be. Here is one:

$$|g(K) - 1| \rightarrow \min \quad (29)$$

$$g(K) \equiv \sqrt{\min(\tilde{\xi}_i^2)} + \sqrt{\max(\tilde{\xi}_i^2)} \quad (30)$$

i.e., we take K for which $g(K)$ (which monotonically decreases with increasing K) is closest to 1. This simple algorithm works pretty well in practical applications.¹⁷

3.2 Limitations

An evident limitation of the principal component approach is that the number of risk factors is limited by M . If long lookbacks are unavailable/undesirable, as, *e.g.*, in short-holding quantitative trading strategies, then typically $M \ll N$. Yet, the number of the actually relevant underlying risk factors can be substantially greater than M , and most of these risk factors are missed by the principal component approach. In this regard, we can ask: can we use other than the first M principal components to build a factor model? The answer, prosaically, is that, without some additional information, it is unclear what to do with the principal components with null eigenvalues. They simply do not contribute to any sample factor covariance matrix. However, not all is lost. There is a way around this difficulty.

4 Heterotic Construction

4.1 Industry Risk Factors

Without long lookbacks, the number of risk factors based on principal components is limited.¹⁸ However, risk factors based on a granular enough industry classification

¹⁷ The distribution of $\tilde{\xi}_i^2$ is skewed; typically, $\tilde{\xi}_i^2$ has a tail at higher values, while $\ln(\tilde{\xi}_i^2)$ has a tail at lower values, and the distribution is only roughly log-normal. So K is not (the floor/cap of) $M/2$, but somewhat higher, albeit close to it. See Table 1 and Figure 1 for an illustrative example.

¹⁸ The number of style factors is also limited (especially for short horizons), of order 10 or fewer.

can be plentiful. Furthermore, they are independent of the pricing data and, in this regard, are insensitive to the lookback. In fact, typically they tend to be rather stable out-of-sample as companies seldom jump industries, let alone sectors.

For terminological definiteness, here we will use the BICS nomenclature for the levels in the industry classification, albeit this is not critical here. Also, BICS has three levels “sector \rightarrow industry \rightarrow sub-industry” (where “sub-industry” is the most granular level). The number of levels in the industry hierarchy is not critical here either. So, we have: N stocks labeled by $i = 1, \dots, N$; K sub-industries labeled by $A = 1, \dots, K$; F industries labeled by $a = 1, \dots, F$; and L sectors labeled by $\alpha = 1, \dots, L$. More generally, we can think of such groupings as “clusters”. Sometimes, loosely, we will refer to such cluster based factors as “industry” factors.¹⁹

4.2 “Binary” Property

The binary property implies that each stock belongs to one and only one sub-industry, industry and sector (or, more generally, cluster). Let G be the map between stocks and sub-industries, S be the map between sub-industries and industries, and W be the map between industries and sectors:

$$G : \{1, \dots, N\} \mapsto \{1, \dots, K\} \quad (31)$$

$$S : \{1, \dots, K\} \mapsto \{1, \dots, F\} \quad (32)$$

$$W : \{1, \dots, F\} \mapsto \{1, \dots, L\} \quad (33)$$

The nice thing about the binary property is that the clusters (sub-industries, industries and sectors) can be used to identify blocks (sub-matrices) in the correlation matrix Ψ_{ij} . *E.g.*, at the most granular level, for sub-industries, the binary matrix $\delta_{G(i),A}$ defines such blocks. Thus, the sum $B_A \equiv \sum_{i=1}^N \delta_{G(i),A} X_i$, where X_i is an arbitrary N -vector, is the same as $\sum_{i \in J(A)} X_i$, where $J(A)$ is the set of tickers in the sub-industry A . These blocks are the backbone of the following construction.

4.3 Heterotic Models

Consider the following factor loadings matrix:

$$\tilde{\Omega}_{iA} = \delta_{G(i),A} U_i \quad (34)$$

$$U_i \equiv [U(A)]_i, \quad i \in J(A), \quad A = 1, \dots, K \quad (35)$$

where $J(A) \equiv \{i | G(i) = A\}$ is the set of tickers (whose number $N(A) \equiv |J(A)|$) in the sub-industry labeled by A . Then the $N(A)$ -vector $U(A)$ is the *first* principal component of the $N(A) \times N(A)$ matrix $\Psi(A)$ defined via $[\Psi(A)]_{ij} \equiv \Psi_{ij}$, $i, j \in J(A)$. (Note that $\sum_{i \in J(A)} [U(A)]_i^2 = 1$; also, let the corresponding (largest) eigenvalue of

¹⁹ Albeit in the BICS context we may be referring to, *e.g.*, sub-industries, while in other classification schemes the actual naming may be altogether different.

$\Psi(A)$ be $\lambda(A)$.²⁰ With this factor loadings matrix we can compute the factor covariance matrix and specific risk via a linear regression as above, and we get:

$$\tilde{\xi}_i^2 = 1 - \lambda(G(i)) U_i^2 \quad (36)$$

$$\left(\tilde{\Omega} \Phi \tilde{\Omega}^T \right)_{ij} = U_i U_j \sum_{k \in J(G(i))} \sum_{l \in J(G(j))} U_k \Psi_{kl} U_l \quad (37)$$

so we have²¹

$$\tilde{\Gamma}_{ij} = [1 - \lambda(G(i)) U_i^2] \delta_{ij} + U_i U_j \sum_{k \in J(G(i))} \sum_{l \in J(G(j))} U_k \Psi_{kl} U_l \quad (38)$$

and automatically $\tilde{\Gamma}_{ii} = 1$. This simplicity is due to the use of the (first) principal components corresponding to the blocks $\Psi(A)$ of the sample correlation matrix.

4.3.1 Multiple Principal Components

For the sake of completeness, let us discuss an evident generalization. Above in (34) we took the binary map between the tickers and sub-industries and augmented it with the first principal components of the corresponding blocks in the sample correlation matrix. Instead of taking only the first principal component, we can take the first $P(A) \geq 1$ principal components for each block labeled by the sub-industry A ($A = 1, \dots, K$). Then we have $\hat{K} = \sum_{A=1}^K P(A)$ risk factors labeled by pairs $\hat{A} \equiv (A, I)$, where for a given value of A we have $I \in D(A)$ (with $|D(A)| = P(A)$). The factor loadings matrix reads:

$$\tilde{\Omega}_{i\hat{A}} = \delta_{G(i),A} [U(A)]_i^{(I)} \quad (39)$$

where $U(A)$ is the $N(A) \times P(A)$ matrix whose columns are the first $P(A)$ principal components (with eigenvalues $[\lambda(A)]^{(I)}$) of the $N(A) \times N(A)$ matrix $\Psi(A)$ (as above, $[\Psi(A)]_{ij} \equiv \Psi_{ij}$, $i, j \in J(A)$, and $\sum_{i \in J(A)} [U(A)]_i^{(I)} [U(A)]_i^{(J)} = \delta_{IJ}$, $I, J \in D(A)$.) In order to have nonvanishing specific risks, it is necessary that we take $P(A) < M$ ($M + 1$ is the number of observations in the time series). We then have

$$\begin{aligned} \Gamma_{ij} = & \left[1 - \sum_{I \in D(G(i))} [\lambda(G(i))]^{(I)} \left([U(G(i))]_i^{(I)} \right)^2 \right] \delta_{ij} + \\ & + \sum_{I \in D(G(i))} \sum_{J \in D(G(j))} [U(G(i))]_i^{(I)} [U(G(j))]_j^{(J)} \times \\ & \times \sum_{k \in J(G(i))} \sum_{l \in J(G(j))} [U(G(i))]_k^{(I)} \Psi_{kl} [U(G(j))]_l^{(J)} \end{aligned} \quad (40)$$

and (as in the case above with all $P(A) \equiv 1$) automatically $\tilde{\Gamma}_{ii} = 1$.

²⁰ If $N(A) = 1$, *i.e.*, we have only one ticker in the sub-industry labeled by A , then $[U(A)]_i = 1$ and $\lambda(A) = \Psi_{ii} = 1$, $i \in J(A)$.

²¹ For single-ticker sub-industries ($N(A) = 1$) the specific risk vanishes: $\tilde{\xi}_i^2 = 0$; however, this does not pose a problem as this does not cause the matrix $\tilde{\Gamma}_{ij}$ to be singular (see below).

4.3.2 Caveats

The above construction might look like a free lunch, but it is not. Let us start with the $P(A) \equiv 1$ case (first principal components only). For short lookbacks, the number of risk factors typically is too large: K can easily be greater than M , so²²

$$\Phi_{AB} = \sum_{i \in J(A)} \sum_{j \in J(B)} U_i \Psi_{ij} U_j \quad (41)$$

is singular. In general, the sample factor covariance matrix is singular if $K > M$. We will deal with this issue below via the nested Russian-doll risk model construction.

This issue is further exacerbated in the multiple principal component construction (with at least some $P(A) > 1$) as the number of risk factors $\hat{K} > K$ is even larger. This too can be dealt with via the Russian-doll construction. However, there is yet another caveat pertinent to using multiple principal components, irrespective of whether the factor covariance matrix is singular or not. The principal components are based on off-diagonal elements of Ψ_{ij} and tend to be unstable out-of-sample, the first principal component typically being the most stable. So, for the sake of simplicity, below we will focus on the case with only first principal components.

5 Russian-Doll Construction

5.1 General Idea

As discussed above, the sample factor covariance matrix Φ_{AB} is singular if the number of factors K is greater than M . The simple idea behind the Russian-doll construction is to model such Φ_{AB} itself via yet another factor model matrix Γ'_{AB} (as opposed to computing it as a sample covariance matrix of the risk factors f_A):²³

$$\Gamma'_{AB} = \sqrt{\Phi_{AA}} \sqrt{\Phi_{BB}} \tilde{\Gamma}'_{AB} \quad (42)$$

$$\tilde{\Gamma}'_{AB} = (\tilde{\xi}'_A)^2 \delta_{AB} + \sum_{a,b=1}^F \tilde{\Omega}'_{Aa} \Phi'_{ab} \tilde{\Omega}'_{Bb} \quad (43)$$

where $\tilde{\xi}'_A$ is the specific risk for the “normalized” factor return $\tilde{f}_A \equiv f_A / \sqrt{\Phi_{AA}}$; $\tilde{\Omega}'_{Aa}$, $A = 1, \dots, K$, $a = 1, \dots, F$ is the corresponding factor loadings matrix; and Φ'_{ab} is the factor covariance matrix for the underlying risk factors f'_a , $a = 1, \dots, F$, where we assume that $F \ll K$. If the smaller factor covariance matrix Φ'_{ab} is still singular, we model it via yet another factor model with fewer risk factors, and so on – until the resulting factor covariance matrix is nonsingular. If, at the final stage, we are left with a single factor, then the resulting 1×1 factor covariance matrix is automatically nonsingular – it is simply the sample variance of the remaining factor.

²² Note that $\Phi_{AA} = \lambda(A)$.

²³ We use a prime on Γ'_{AB} , $\tilde{\xi}'_A$, Φ'_{ab} , *etc.* to avoid confusion with Γ_{ij} , $\tilde{\xi}_i$, Φ_{AB} , *etc.*

5.2 Complete Heterotic Russian-Doll Embedding

For concreteness we will use the BICS terminology for the levels in the industry classification, albeit this is not critical here. Also, BICS has three levels “sector \rightarrow industry \rightarrow sub-industry” (where “sub-industry” is the most granular level). For definiteness, we will assume three levels here, and the generalization to more levels is straightforward. So, we have: N stocks labeled by $i = 1, \dots, N$; K sub-industries labeled by $A = 1, \dots, K$; F industries labeled by $a = 1, \dots, F$; and L sectors labeled by $\alpha = 1, \dots, L$. A nested Russian-doll risk model then is constructed as follows:

$$\Gamma_{ij} = \sqrt{C_{ii}} \sqrt{C_{jj}} \tilde{\Gamma}_{ij} \quad (44)$$

$$\tilde{\Gamma}_{ij} = \tilde{\xi}_i^2 \delta_{ij} + U_i U_j \Gamma'_{G(i), G(j)} \quad (45)$$

$$\Gamma'_{AB} = \sqrt{\Phi_{AA}} \sqrt{\Phi_{BB}} \tilde{\Gamma}'_{AB} \quad (46)$$

$$\tilde{\Gamma}'_{AB} = (\tilde{\xi}'_A)^2 \delta_{AB} + U'_A U'_B \Gamma''_{S(A), S(B)} \quad (47)$$

$$\Gamma''_{ab} = \sqrt{\Phi'_{aa}} \sqrt{\Phi'_{bb}} \tilde{\Gamma}''_{ab} \quad (48)$$

$$\tilde{\Gamma}''_{ab} = (\tilde{\xi}''_a)^2 \delta_{ab} + U''_a U''_b \Gamma'''_{W(a), W(b)} \quad (49)$$

$$\Gamma'''_{\alpha\beta} = \sqrt{\Phi''_{\alpha\alpha}} \sqrt{\Phi''_{\beta\beta}} \tilde{\Gamma}'''_{\alpha\beta} \quad (50)$$

$$\tilde{\Gamma}'''_{\alpha\beta} = (\tilde{\xi}'''_\alpha)^2 \delta_{\alpha\beta} + U'''_\alpha U'''_\beta \Phi''' \quad (51)$$

where

$$\tilde{\xi}_i^2 = 1 - \lambda(G(i)) U_i^2 \quad (52)$$

$$(\tilde{\xi}'_A)^2 = 1 - \lambda'(S(A)) (U'_A)^2 \quad (53)$$

$$(\tilde{\xi}''_a)^2 = 1 - \lambda''(W(a)) (U''_a)^2 \quad (54)$$

$$(\tilde{\xi}'''_\alpha)^2 = 1 - \lambda'''(U'''_\alpha)^2 \quad (55)$$

and

$$\Phi_{AB} = \sum_{i \in J(A)} \sum_{j \in J(B)} U_i \Psi_{ij} U_j \quad (56)$$

$$\Phi'_{ab} = \sum_{A \in J'(a)} \sum_{B \in J'(b)} U'_A \Psi'_{AB} U'_B \quad (57)$$

$$\Phi''_{\alpha\beta} = \sum_{a \in J''(\alpha)} \sum_{b \in J''(\beta)} U''_a \Psi''_{ab} U''_b \quad (58)$$

$$\Phi''' = \sum_{\alpha, \beta=1}^L U'''_\alpha \Psi'''_{\alpha\beta} U'''_\beta \quad (59)$$

so we have $\Phi_{AA} = \lambda(A)$, $\Phi'_{aa} = \lambda'(a)$, $\Phi''_{\alpha\alpha} = \lambda''(\alpha)$, and $\Phi''' = \lambda'''$ (see below). Also, $J(A) = \{i | G(i) = A\}$ ($N_A \equiv |J(A)|$ tickers in sub-industry A), $J'(a) =$

$\{A|S(A) = a\}$ ($N'(a) \equiv |J'(a)|$ sub-industries in industry a), $J''(\alpha) = \{a|W(a) = \alpha\}$ ($N''(\alpha) \equiv |J''(\alpha)|$ industries in sector α), and the maps G (tickers to sub-industries), S (sub-industries to industries) and W (industries to sectors) are defined in (31), (32) and (33). Furthermore,

$$\Psi_{ij} = C_{ij} / \sqrt{C_{ii}} \sqrt{C_{jj}} \quad (60)$$

$$\Psi'_{AB} = \Phi_{AB} / \sqrt{\Phi_{AA}} \sqrt{\Phi_{BB}} \quad (61)$$

$$\Psi''_{ab} = \Phi'_{ab} / \sqrt{\Phi'_{aa}} \sqrt{\Phi'_{bb}} \quad (62)$$

$$\Psi'''_{\alpha\beta} = \Phi''_{\alpha\beta} / \sqrt{\Psi''_{\alpha\alpha}} \sqrt{\Psi''_{\beta\beta}} \quad (63)$$

and

$$U_i \equiv [U(A)]_i, \quad i \in J(A) \quad (64)$$

$$U'_A \equiv [U'(a)]_A, \quad A \in J'(a) \quad (65)$$

$$U''_a \equiv [U''(\alpha)]_a, \quad a \in J''(\alpha) \quad (66)$$

The $N(A)$ -vector $U(A)$ is the first principal component of $\Psi(A)$ with the eigenvalue $\lambda(A)$ ($[\Psi(A)]_{ij} \equiv \Psi_{ij}$, $i, j \in J(A)$), the $N'(a)$ -vector $U'(a)$ is the first principal component of $\Psi'(a)$ with the eigenvalue $\lambda'(a)$ ($[\Psi'(a)]_{AB} \equiv \Psi'_{AB}$, $A, B \in J'(a)$), the $N''(\alpha)$ -vector $U''(\alpha)$ is the first principal component of $\Psi''(\alpha)$ with the eigenvalue $\lambda''(\alpha)$ ($[\Psi''(\alpha)]_{ab} \equiv \Psi''_{ab}$, $a, b \in J''(\alpha)$), while U'''_{α} is the first principal component of $\Psi'''_{\alpha\beta}$ with the eigenvalue λ''' . The vectors $U(A)$, $U'(a)$ and $U''(\alpha)$ are normalized, so $\sum_{i \in J(A)} U_i^2 = 1$, $\sum_{A \in J'(a)} (U'_A)^2 = 1$, $\sum_{a \in J''(\alpha)} (U''_a)^2 = 1$, and also $\sum_{\alpha=1}^L (U'''_{\alpha})^2 = 1$.

For the sake of completeness, above we included the step where the sample factor covariance matrix $\Phi''_{\alpha\beta}$ for the sectors is further approximated via a 1-factor model $\tilde{\Gamma}'''_{\alpha\beta}$. If $\Phi''_{\alpha\beta}$ computed via (58) is nonsingular, then this last step can be omitted,²⁴ so at the last stage we have L factors (as opposed to a single factor).²⁵ Similarly, if we have enough observations to compute the sample covariance matrix Φ'_{ab} for the industries, we can stop at that stage. Finally, note that in the above construction we are guaranteed to have $(\tilde{\xi}'''_{\alpha})^2 > 0$, $(\tilde{\xi}''_{\alpha})^2 > 0$, $(\tilde{\xi}'_A)^2 > 0$ and $\tilde{\xi}_i^2 \geq 0$ (with the last equality occurring only for single-ticker sub-industries and not posing a problem – see below).²⁶ In Appendix B we give the R code for building heterotic risk models.

²⁴ That is, assuming there are enough observations in the time series for out-of-sample stability.

²⁵ This last factor can be interpreted as the “market” risk factor. For the sake of completeness, the definitions of the factors at each stage are as follows: (i) for the sub-industries $f_A = \sum_{i \in J(A)} U_i \tilde{R}_i$, where $\tilde{R}_i = R_i / \sqrt{C_{ii}}$; (ii) for the industries $f'_a = \sum_{A \in J'(a)} U'_A \tilde{f}_A$, where $\tilde{f}_A = f_A / \sqrt{\Phi_{AA}}$; (iii) for the sectors $f''_{\alpha} = \sum_{a \in J''(\alpha)} U''_a \tilde{f}'_a$, where $\tilde{f}'_a = f'_a / \sqrt{\Phi'_{aa}}$; and (iv) for the “market” $f''' = \sum_{\alpha=1}^L U'''_{\alpha} \tilde{f}''_{\alpha}$, where $\tilde{f}''_{\alpha} = f''_{\alpha} / \sqrt{\Phi''_{\alpha\alpha}}$.

²⁶ For a typical, large trading universe, industries and sectors usually contain more than one ticker; however, there can be cases of single-ticker sub-industries. Nonetheless, $\tilde{\Gamma}_{ij}$ is nonsingular. Indeed, for an arbitrary N -vector X_i we have $X^T \tilde{\Gamma} X > 0$ unless $X_i \equiv 0$, $i \notin H$, where $H \equiv \{i | N(G(i)) = 1\}$. For such X_i we have $X^T \tilde{\Gamma} X = \sum_{A, B \in E} Y_A \Gamma'_{AB} Y_B > 0$, where $E \equiv \{A | N(A) = 1\}$, $Y_A \equiv \delta_{A, G(i)} X_i$, $A \in E$, and we have taken into account that by construction Γ'_{AB} (and its sub-matrix with $A, B \in E$) is positive-definite, and also that $U_i = 1$, $i \in H$. More on this below.

5.3 Model Covariance Matrix and Its Inverse

The model covariance matrix is given by Γ_{ij} defined in (44). For completeness, let us present it in the “canonical” form:

$$\Gamma_{ij} = \xi_i^2 \delta_{ij} + \sum_{A,B=1}^N \Omega_{iA} \Phi_{AB}^* \Omega_{jB} \quad (67)$$

where

$$\xi_i^2 \equiv \sqrt{C_{ii}} \sqrt{C_{jj}} \tilde{\xi}_i^2 \quad (68)$$

$$\Omega_{iA} \equiv \sqrt{C_{ii}} U_i \delta_{G(i),A} \quad (69)$$

$$\Phi_{AB}^* \equiv \Gamma'_{AB} \quad (70)$$

where $\tilde{\xi}_i^2$ is defined in (52), U_i is defined in (64), Γ'_{AB} is defined in (46), and we use the star superscript in the our factor covariance matrix Φ_{AB}^* (which is nonsingular) to distinguish it from the sample factor covariance matrix Φ_{AB} (which is singular).

In many applications, such as portfolio optimization, one needs the inverse of the matrix Γ . When we have no single-ticker sub-industries, the inverse is given by (in matrix notation)

$$\Gamma^{-1} = \Xi^{-1} - \Xi^{-1} \Omega \Delta^{-1} \Omega^T \Xi^{-1} \quad (71)$$

$$\Delta \equiv (\Phi^*)^{-1} + \Omega^T \Xi^{-1} \Omega \quad (72)$$

$$\Xi \equiv \text{diag}(\xi_i^2) \quad (73)$$

However, when there are some single-ticker sub-industries, the corresponding $\xi_i^2 = 0$, $i \in H$ ($H \equiv \{i | N(G(i)) = 1\}$), so (71) “breaks”. Happily, there is an easy “fix”. This is because for such tickers the specific risk and factor risk are *indistinguishable*. Recall that $U_i = 1$, $i \in H$, and $\Phi_{AA}^* = 1$, $A \in E$ ($E \equiv \{A | G(A) = 1\}$). We can rewrite Γ_{ij} via

$$\Gamma_{ij} = \hat{\xi}_i^2 \delta_{ij} + \sum_{A,B=1}^N \Omega_{iA} \hat{\Phi}_{AB}^* \Omega_{jB} \quad (74)$$

where: $\hat{\xi}_i^2 = \xi_i^2$ for $i \notin H$; $\hat{\xi}_i^2 = C_{ii} \zeta_i$ for $i \in H$ with arbitrary ζ_i , $0 < \zeta_i \leq 1$; $\hat{\Phi}_{AB}^* = \Phi_{AB}^*$ if $A \notin E$ or $B \notin E$ or $A \neq B$; and $\hat{\Phi}_{AA}^* = 1 - \delta_{A,G(i)} \zeta_i$ for $A \in E$. (Here we have taken into account that $U_i = 1$, $i \in H$.) Now we can invert Γ via

$$\Gamma^{-1} = \hat{\Xi}^{-1} - \hat{\Xi}^{-1} \Omega \hat{\Delta}^{-1} \Omega^T \hat{\Xi}^{-1} \quad (75)$$

$$\hat{\Delta} \equiv (\hat{\Phi}^*)^{-1} + \Omega^T \hat{\Xi}^{-1} \Omega \quad (76)$$

$$\hat{\Xi} \equiv \text{diag}(\hat{\xi}_i^2) \quad (77)$$

Note that, due to the factor model structure, to invert the $N \times N$ matrix Γ , we only need to invert two $K \times K$ matrices $\hat{\Phi}^*$ and $\hat{\Delta}$. If there are no single-ticker sub-industries, then Φ^* itself has a factor model structure and involves inverting two $F \times F$ matrices, one of which has a factor model structure, and so on.

6 Horse Race

So, suppose we have built a complete heterotic risk model. How do we know it adds value? *I.e.*, how do we know that the off-diagonal elements of the factor model covariance matrix Γ_{ij} are stable out-of-sample to the extent that they add value. We can run a horse race. There are many ways of doing this. Here is one. For a given trading universe we compute some expected returns, *e.g.*, based on overnight mean-reversion. We can construct a trading portfolio by using our heterotic risk model covariance matrix in the optimization whereby we maximize the Sharpe ratio (subject to the dollar neutrality constraint). On the other hand, we can run the same optimization with a diagonal sample covariance matrix $\text{diag}(C_{ii})$ subject to neutrality (via linear homogeneous constraints) w.r.t. the underlying heterotic risk factors (plus dollar neutrality).²⁷ In fact, optimization with such diagonal covariance matrix and subject to such linear homogeneous constraints is equivalent to a weighted cross-sectional regression with the loadings matrix (over which the expected returns are regressed) identified with the factor loadings matrix (augmented by the intercept, *i.e.*, the unit vector, for dollar neutrality) and the regression weights identified with the inverse sample variances $1/C_{ii}$ (see (Kakushadze, 2015) for details). So, we will refer to the horse race as between optimization (using the heterotic risk model) and weighted regression (with the aforementioned linear homogeneous constraints).²⁸

6.1 Notations

Let P_{is} be the time series of stock prices, where $i = 1, \dots, N$ labels the stocks, and $s = 0, 1, \dots, M$ labels the trading dates, with $s = 0$ corresponding to the most recent date in the time series. The superscripts O and C (unadjusted open and close prices) and AO and AC (open and close prices fully adjusted for splits and dividends) will distinguish the corresponding prices, so, *e.g.*, P_{is}^C is the unadjusted close price. V_{is} is the unadjusted daily volume (in shares). Also, for each date s we define the overnight return as the previous-close-to-open return:

$$E_{is} \equiv \ln(P_{is}^{AO}/P_{i,s+1}^{AC}) \quad (78)$$

This return will be used in the definition of the expected return in our mean-reversion alpha. We will also need the close-to-close return

$$R_{is} \equiv \ln(P_{is}^{AC}/P_{i,s+1}^{AC}) \quad (79)$$

An out-of-sample (see below) time series of these returns will be used in constructing the heterotic risk model and computing, among other things, the sample variances C_{ii} . Also note that all prices in the definitions of E_{is} and R_{is} are fully adjusted.

²⁷ For comparative purposes, we will also run separate backtests where we require neutrality w.r.t. BICS sub-industries and principal components.

²⁸ The remainder of this section somewhat overlaps with Section 7 of (Kakushadze, 2015a) as the backtesting models are similar, albeit not identical.

We assume that: i) the portfolio is established at the open²⁹ with fills at the open prices P_{is}^O ; ii) it is liquidated at the close on the same day – so this is a purely intraday alpha – with fills at the close prices P_{is}^C ; and iii) there are no transaction costs or slippage – our aim here is not to build a realistic trading strategy, but to test that our heterotic risk model adds value to the alpha. The P&L for each stock

$$\Pi_{is} = H_{is} \left[\frac{P_{is}^C}{P_{is}^O} - 1 \right] \quad (80)$$

where H_{is} are the *dollar* holdings. The shares bought plus sold (establishing plus liquidating trades) for each stock on each day are computed via $Q_{is} = 2|H_{is}|/P_{is}^O$.

6.2 Universe Selection

For the sake of simplicity,³⁰ we select our universe based on the average daily dollar volume (ADDV) defined via (note that A_{is} is out-of-sample for each date s):

$$A_{is} \equiv \frac{1}{d} \sum_{r=1}^d V_{i,s+r} P_{i,s+r}^C \quad (81)$$

We take $d = 21$ (*i.e.*, one month), and then take our universe to be the top 2000 tickers by ADDV. To ensure that we do not inadvertently introduce a universe selection bias, we rebalance monthly (every 21 trading days, to be precise). *I.e.*, we break our 5-year backtest period (see below) into 21-day intervals, we compute the universe using ADDV (which, in turn, is computed based on the 21-day period immediately preceding such interval), and use this universe during the entire such interval. We do have the survivorship bias as we take the data for the universe of tickers as of 9/6/2014 that have historical pricing data on <http://finance.yahoo.com> (accessed on 9/6/2014) for the period 8/1/2008 through 9/5/2014. We restrict this universe to include only U.S. listed common stocks and class shares (no OTCs, preferred shares, *etc.*) with BICS sector, industry and sub-industry assignments as of 9/6/2014.³¹ However, as discussed in detail in Section 7 of (Kakushadze, 2015a), the survivorship bias is not a leading effect in such backtests.³²

6.3 Backtesting

We run our simulations over a period of 5 years (more precisely, 1260 trading days going back from 9/5/2014, inclusive). The annualized return-on-capital (ROC) is

²⁹ This is a so-called “delay-0” alpha: the same price, P_{is}^O (or adjusted P_{is}^{AO}), is used in computing the expected return (via E_{is}) and as the establishing fill price.

³⁰ In practical applications, the trading universe of liquid stocks typically is selected based on market cap, liquidity (ADDV), price and other (proprietary) criteria.

³¹ The choice of the backtesting window is based on what data was readily available.

³² Here we are after the *relative outperformance*, and it is reasonable to assume that, to the leading order, individual performances are affected by the survivorship bias approximately equally.

computed as the average daily P&L divided by the intraday investment level I (with no leverage) and multiplied by 252. The annualized Sharpe Ratio (SR) is computed as the daily Sharpe ratio multiplied by $\sqrt{252}$. Cents-per-share (CPS) is computed as the total P&L divided by the total shares traded.³³

6.4 Weighted Regression Alphas

We will always require that our portfolio be dollar neutral:

$$\sum_{i=1}^N H_{is} = 0 \quad (82)$$

We will further require neutrality

$$\sum_{i=1}^N H_{is} \Lambda_{iAs} = 0 \quad (83)$$

with the three different incarnations for the loadings matrix Λ_{iA} (for each trading day s , so we omit the index s)³⁴ defined via:

$$\text{principal components: } \Lambda_{iA} = \sqrt{C_{ii}} \sqrt{\lambda^{(A)}} V_i^{(A)}, \quad A = 1, \dots, K_{PC} \quad (84)$$

$$\text{sub-industries: } \Lambda_{iA} = \delta_{G(i),A}, \quad A = 1, \dots, K \quad (85)$$

$$\text{heterotic risk factors: } \Lambda_{iA} = \sqrt{C_{ii}} U_i \delta_{G(i),A}, \quad A = 1, \dots, K \quad (86)$$

Here $V_i^{(A)}$ are the first K_{PC} principal components (with the eigenvalues $\lambda^{(A)}$)³⁵ of the sample correlation matrix Ψ_{ij} . For each date s we take $M+1 = d = 21$ trading days as our lookback (*i.e.*, the number of observations) in the out-of-sample time series of close-to-close (see (79)) returns $(R_{i,(s+1)}, R_{i,(s+2)}, \dots, R_{i,(s+d)})$ (based on which we compute the sample covariance (correlation) matrix C_{ijs} (Ψ_{ijs}) for each s), so the number of the nonvanishing eigenvalues $\lambda^{(A)} > 0$ is $M = 20$, and we take $K_{PC} = M$. Further, the map G between tickers and sub-industries is defined in (31), and K is the number of sub-industries.³⁶ Finally, the vector U_i in (86) is defined in (64).

For each date labeled by s , we run cross-sectional regressions of the overnight (see (78)) returns E_{is} over the corresponding loadings matrix, call it Y (with indices

³³ As mentioned above, we assume no transaction costs, which are expected to reduce the ROC of the optimization and weighted regression alphas by the same amount as the two strategies trade the exact same amount by design. Therefore, including the transaction costs would have no effect on the actual *relative outperformance* in the horse race, which is what we are after here.

³⁴ The loadings Λ_{iA} in (84) and (86) are computed for each trading date s (as opposed to, say, every 21 days – see below); in (85) they change only with the universe (every 21 days).

³⁵ The factor $\sqrt{\lambda^{(A)}}$ in (84) does not affect the regression residuals below.

³⁶ In (85) we deliberately take $\Lambda_{iA} = \delta_{G(i),A}$ as opposed to $\Lambda_{iA} = \sqrt{C_{ii}} \delta_{G(i),A}$ (see below). Note that with (85) the intercept is subsumed in Λ_{iA} as we have $\sum_{A=1}^K \Lambda_{iA} = 1$, so (82) is automatic.

suppressed), which has 3 different incarnations: i) for principal components, Y is an $N \times (K_{PC} + 1)$ matrix, whose first column is the intercept (unit N -vector), and the remaining columns are populated by Λ_{iA} defined in (84); ii) for sub-industries, the elements of Y are the same as Λ_{iA} defined in (85); and iii) for heterotic risk factors, Y is an $N \times (K + 1)$ matrix, whose first column is the intercept (unit N -vector), and the remaining columns are populated by Λ_{iA} defined in (86). We take the regression weights to be $z_i \equiv 1/C_{ii}$. More precisely, to avoid unnecessary variations in the weights z_i (as such variations could result in unnecessary overtrading), we do not recompute z_i daily but every 21 trading days, same as with the trading universe.

In the cases i)-iii) above, we compute the residuals ε_{is} of the weighted regression and the dollar holdings H_{is} via (we use matrix notation and suppress indices):

$$\tilde{E} \equiv Z \varepsilon = Z [E - Y (Y^T Z Y)^{-1} Y^T Z E] \quad (87)$$

$$H_{is} \equiv -\tilde{E}_{is} \frac{I}{\sum_{j=1}^N |\tilde{E}_{js}|} \quad (88)$$

where $Z \equiv \text{diag}(z_i)$, we have dollar neutrality (82),³⁷ and $\sum_{i=1}^N |H_{is}| = I$ (the total *intraday* dollar investment level (long plus short), which is the same for all dates s).

The simulation results are given in Table 2 and P&Ls for the 3 cases i)-iii) are plotted in Figure 2. For comparison purposes – and to allay any potential concerns that the results in Table 2 may not hold for realistic position bounds, in Table 3 we give the simulation results for the same cases i)-iii) above with the strict bounds

$$|H_{is}| \leq 0.01 A_{is} \quad (89)$$

so not more than 1% of each stock's ADDV is bought or sold. We use the bounded regression algorithm and the R source code of (Kakushadze, 2015b) to run these simulations. Expectedly, the liquidity bounds (89) lower ROC and CPS while improving SR, but in the same fashion for all 3 weighted regression alphas i)-iii). The results in Tables 2 and 3 confirm our prior intuitive argument that the sub-industries outperform the principal components simply because they are more numerous.³⁸

If we compute Λ_{iA} in (84) and (86) every 21 trading days (instead of daily – see fn. 34), the difference is very slight. *E.g.*, for the heterotic risk factors computed every 21 days (with no bounds) we get: ROC = 51.66%, SR = 13.42, CPS = 2.26.

Finally, let us also mention that, in the weighted regressions ii) and iii), the dollar holdings for the tickers in the single-ticker sub-industries are automatically null. This is *not* the case for optimized alphas (see below). Generally, if single-ticker (or small) sub-industries are undesirable, one can “prune” the industry hierarchy tree by merging (single-ticker and/or small) sub-industries at the industry level.

³⁷ Due to \tilde{E}_{is} having 0 cross-sectional means, which in turn is due to the intercept either being included (the cases i) and iii)), or being subsumed in the loadings matrix Y (the case ii)).

³⁸ In Table 2 the heterotic risk factors outperform the sub-industries. However, this is largely an artifact of defining Λ_{iA} as in (85). If we take $\Lambda_{iA} = \sqrt{C_{ii}} \delta_{G(i),A}$ instead (and augment the regression loadings matrix Y with the intercept for dollar neutrality), we will get (without the bounds (89) – the results with the bounds are similar): ROC = 51.62%, SR = 13.45, CPS = 2.26.

6.5 Optimized Alphas

As mentioned above, our goal is to determine whether the heterotic risk model construction adds value by comparing the simulated performance of the weighted regression alphas above to the simulated performance of the optimized alphas (via maximizing the Sharpe ratio) based on the same expected returns E_{is} . In maximizing the Sharpe ratio, we use the heterotic risk model covariance matrix Γ_{ij} given by (44), which we compute every 21 trading days (same as for the universe). For each date (we omit the index s) we maximize the Sharpe ratio subject to the dollar neutrality constraint:

$$\mathcal{S} \equiv \frac{\sum_{i=1}^N H_i E_i}{\sqrt{\sum_{i,j=1}^N \Gamma_{ij} H_i H_j}} \rightarrow \max \quad (90)$$

$$\sum_{i=1}^N H_i = 0 \quad (91)$$

The solution is given by

$$H_i = -\gamma \left[\sum_{j=1}^N \Gamma_{ij}^{-1} E_j - \sum_{j=1}^N \Gamma_{ij}^{-1} \frac{\sum_{k,l=1}^N \Gamma_{kl}^{-1} E_l}{\sum_{k,l=1}^N \Gamma_{kl}^{-1}} \right] \quad (92)$$

where Γ^{-1} is the inverse of Γ (see Subsection 5.3), and the overall normalization constant $\gamma > 0$ (this is a mean-reversion alpha) is fixed via the requirement that

$$\sum_{i=1}^N |H_i| = I \quad (93)$$

Note that (92) satisfies the dollar neutrality constraint (91).

The simulation results are given in Table 2 in the bottom row. The P&L plot for this optimized alpha is included in Figure 2. For the same reasons as in the case of weighted regression alphas, in the bottom row of Table 3 we give the simulation results for the same optimized alpha with the strict liquidity bounds (89).³⁹ We use the optimization algorithm for maximizing the Sharpe ratio subject to linear homogeneous constraints and bounds discussed in (Kakushadze, 2015a).⁴⁰ Also, in the second rows in Tables 2 and 3 we have included the simulation results for the optimized alpha where in the optimization we use the risk factor model covariance matrix Γ_{ij} based on the principal components discussed in Section 3.⁴¹ From our

³⁹ In Tables 2 and 3 at the final stage the heterotic risk factors are the (10) BICS sectors: there are enough (20) observations in the time series. The 1-factor model gives almost the same results.

⁴⁰ The source code for this algorithm is not included in (Kakushadze, 2015a), so we include it in Appendix C. It is similar to the source code of (Kakushadze, 2015b) for the bounded regression.

⁴¹ This matrix is given by $\Gamma_{ij} = \sqrt{C_{ii}}\sqrt{C_{jj}}\tilde{\Gamma}_{ij}$, where $\tilde{\Gamma}_{ij}$ is defined in (27), and K is determined via the algorithm of Section 3.1. For the $d = 21$ trading day lookback in our backtests, the value of K fixed by this algorithm turns out to be $K = 13$.

simulation results in Tables 2 and 3 it is evident that the heterotic risk model predicts off-diagonal elements of the covariance matrix (that is, correlations) out-of-sample rather well. Indeed, using it in the optimization sizably improves ROC, SR and CPS compared with the weighted regressions with all three loadings i)-iii) above.⁴²

7 Concluding Remarks

The heterotic risk model construction we discuss in this paper is based on a “heterosis” of: i) granularity of an industry classification; ii) diagonality of the principal component factor covariance matrix for any sub-cluster of stocks; and iii) dramatic reduction of the size of the factor covariance matrix in the Russian-doll construction. This is a powerful approach, as is evident from the horse race we ran above.

Naturally, one may wonder if we can extend our construction to risk models which do not include any statistical risk factors (*i.e.*, principal components) or include other non-binary factors such as style factors. A key simplifying feature in the heterotic construction is that the industry classification, which is used as the backbone (and is augmented with the principal components to satisfy the conditions (8)), is binary. Once non-binary risk factors are included, it is more nontrivial to compute the specific risk and the factor covariance matrix (such that (8) are satisfied). However, there exist proprietary algorithms for dealing with this, which are outside of the scope of this paper. We hope to make these algorithms a public knowledge elsewhere.

One final remark concerns purely statistical risk models based on principal components. Albeit their market share is rather limited, it is unclear why a portfolio manager would be willing to pay for such models considering that they are straightforward to build in-house, especially now that we have provided the source code for constructing them. One argument is that using option implied volatility (which is available only for optionable stocks) to model stock volatility should work better,⁴³ and if a portfolio manager does not possess the implied volatility data or the know-how for incorporating it into a statistical risk model, he or she would be better off simply buying one from a provider. However, this argument appears to be thin, at best. Nowadays, with ever-shortening lookbacks, it is unclear if the implied volatility indeed adds any value when the risk model is used in actual portfolio optimization for actual alphas. In this regard a new study would appear to be warranted. In any event, as we saw above, heterotic risk models outperform principal component risk models by a significant margin, so one can build heterotic risk models in-house (instead of buying less powerful statistical models) now that this know-how is in the public domain. The only data needed to construct a heterotic risk model is: i) adjusted close prices; and ii) a granular enough binary industry classification, such as GICS, BICS, ICB, *etc.* Most quantitative traders already have this data in-house. So, we hope this paper further encourages/aids organic custom risk model building.

⁴² There exist further (proprietary) performance improvements using the heterotic risk model.

⁴³ In this context, the paper (Ederington and Guan, 2002) sometimes is referred to.

A R Code: Principal Component Risk Model

In this appendix we give the R (R Package for Statistical Computing, <http://www.r-project.org>) source code for building a purely statistical risk model (principal components) based on the algorithm we discuss in Section 3, including the algorithm for fixing the number of factors K in Section 3.1. The code below is essentially self-explanatory and straightforward. It consists of a single function `qrm.cov.pc(ret, use.cor = T)`. The input is: i) `ret`, an $N \times d$ matrix of returns (*e.g.*, daily close-to-close returns), where N is the number of tickers, d is the number of observations in the time series (*e.g.*, the number of trading days), and the ordering of the dates is immaterial; and ii) `use.cor`, where for `TRUE` (default) the risk factors are computed based on the principal components of the sample correlation matrix Ψ_{ij} , whereas for `FALSE` they are computed based on the sample covariance matrix C_{ij} . The output is a list: `result$spec.risk` is the specific risk ξ_i (not the specific variance ξ_i^2), `result$fac.load` is the factor loadings matrix $\Omega_{iA} = \sqrt{C_{ii}} \tilde{\Omega}_{iA}$, `result$fac.cov` is the factor covariance matrix Φ_{AB} (with the normalization (26) for the factor loadings matrix, $\Phi_{AB} = \delta_{AB}$), `result$cov.mat` is the factor model covariance matrix $\Gamma_{ij} = \sqrt{C_{ii}}\sqrt{C_{jj}} \tilde{\Gamma}_{ij}$, and `result$inv.cov` is the matrix Γ_{ij}^{-1} inverse to Γ_{ij} .

```
qrm.cov.pc <- function (ret, use.cor = T)
{
  print("Running qrm.cov.pc()...")

  tr <- apply(ret, 1, sd)
  if(use.cor)
    ret <- ret / tr

  d <- ncol(ret)
  x <- t(ret)
  x <- var(x, x)
  tv <- diag(x)
  x <- eigen(x)

  g.prev <- 999
  for(k in 1:(d-1))
  {
    u <- x$values[1:k]
    v <- x$vectors[, 1:k]
    v <- t(sqrt(u) * t(v))
    x.f <- v %*% t(v)
    x.s <- tv - diag(x.f)
    z <- x.s / tv
    g <- abs(sqrt(min(z)) + sqrt(max(z)) - 1)
```

```

    if(g > g.prev)
      break

    g.prev <- g

    spec.risk <- sqrt(x.s)
    fac.load <- v
    fac.cov <- diag(1, k)
    cov.mat <- diag(x.s) + x.f
  }

  k <- k - 1
  y.s <- 1 / spec.risk / spec.risk
  v <- fac.load
  v1 <- y.s * v
  inv.cov <- diag(y.s) - v1 %*% solve(diag(1, k) + t(v) %*% v1) %*% t(v1)

  if(use.cor)
  {
    spec.risk <- tr * spec.risk
    fac.load <- tr * fac.load
    cov.mat <- tr * t(tr * cov.mat)
    inv.cov <- t(inv.cov / tr) / tr
  }

  result <- new.env()
  result$spec.risk <- spec.risk
  result$fac.load <- fac.load
  result$fac.cov <- fac.cov
  result$cov.mat <- cov.mat
  result$inv.cov <- inv.cov
  result <- as.list(result)
  return(result)
}

```

B R Code: Heterotic Risk Model

In this appendix we give the R source code for building the heterotic risk model based on the algorithm we discuss in Section 5.2. The code below is essentially self-explanatory and straightforward as it simply follows the formulas in Section 5.2. It consists of a single function `qrm.het(ret, ind, mkt.fac = F, rm.sing.tkr = F)`.

The input is: i) `ret`, an $N \times d$ matrix of returns (*e.g.*, daily close-to-close returns), where N is the number of tickers, d is the number of observations in the time series (*e.g.*, the number of trading days), and the ordering of the dates is immaterial; ii) `ind` is a list whose length *a priori* is arbitrary, and its elements are populated by the binary matrices (with rows corresponding to tickers, so `dim(ind[[1]])[1]` is N) corresponding to the levels in the input binary industry classification hierarchy in the order of decreasing granularity (so, in the BICS case `ind[[1]]` is the $N \times K$ matrix $\delta_{G(i),A}$ (sub-industries), `ind[[2]]` is the $N \times F$ matrix $\delta_{G'(i),a}$ (industries), and `ind[[3]]` is the $N \times L$ matrix $\delta_{G''(i),\alpha}$ (sectors), where the map G is defined in (31) (tickers to sub-industries), $G' \equiv GS$ (tickers to industries), and $G'' \equiv GSW$ (tickers to sectors), with the map S (sub-industries to industries) defined in (32), and the map W (industries to sectors) defined in (33)); iii) `mkt.fac`, where for `TRUE` at the final step we have a single factor (“market”), while for `FALSE` (default) the factors correspond to the least granular level in the industry classification hierarchy; and iv) `rm.sing.tkr`, where for `TRUE` the tickers corresponding to the single-ticker clusters at the most granular level in the industry classification hierarchy (in the BICS case this would be the sub-industry level) are dropped altogether, while for `FALSE` (default) the output universe is the same as the input universe. The output is a list: `result$spec.risk` is the specific risk ξ_i (not the specific variance ξ_i^2), `result$fac.load` is the factor loadings matrix $\Omega_{iA} = \sqrt{C_{ii}} \tilde{\Omega}_{iA}$, `result$fac.cov` is the factor covariance matrix Φ_{AB} , `result$cov.mat` is the factor model covariance matrix $\Gamma_{ij} = \sqrt{C_{ii}}\sqrt{C_{jj}} \tilde{\Gamma}_{ij}$, and `result$inv.cov` is the matrix Γ_{ij}^{-1} inverse to Γ_{ij} .

```
qrm.het <- function (ret, ind, mkt.fac = F, rm.sing.tkr = F)
{
  print("Running qrm.het()...")

  if(rm.sing.tkr)
  {
    bad <- colSums(ind[[1]]) == 1
    ind[[1]] <- ind[[1]][, !bad]
    bad <- rowSums(ind[[1]]) == 0
    for(lvl in 1:length(ind))
      ind[[lvl]] <- ind[[lvl]][!bad, ]
    ret <- ret[!bad, ]
  }

  cov.mat <- list()
  u <- list()
  flm <- ind

  calc.load <- function(load, load1)
  {
```

```

    x <- colSums(load1)
    load <- (t(load1) %*% load) / x
    return(load)
}

calc.cor.mat <- function(cov.mat)
{
  tr <- sqrt(diag(cov.mat))
  cor.mat <- t(cov.mat / tr) / tr
  return(cor.mat)
}

calc.cov.mat <- function(cor.mat, tr)
{
  cov.mat <- t(cor.mat * tr) * tr
  return(cov.mat)
}

cov.mat[[1]] <- var(t(ret), t(ret))
cor.mat <- calc.cor.mat(cov.mat[[1]])

for(lvl in 1:length(ind))
{
  if(lvl > 1)
    flm[[lvl]] <- calc.load(ind[[lvl]], ind[[lvl-1]])

  u[[lvl]] <- rep(NA, nrow(flm[[lvl]]))

  for(a in 1:ncol(flm[[lvl]]))
  {
    take <- as.logical(flm[[lvl]][, a])
    x <- cor.mat[take, take]
    y <- eigen(x)$vectors
    y <- y[, 1]
    u[[lvl]][take] <- y
  }

  flm[[lvl]] <- u[[lvl]] * flm[[lvl]]

  cm <- cov.mat[[lvl + 1]] <- t(flm[[lvl]]) %*% cor.mat %*% flm[[lvl]]
  cor.mat <- calc.cor.mat(cm)
}

```



```

if(!mkt.fac)
  mod.mat <- cm
else
{
  k <- nrow(cor.mat)
  x <- eigen(cor.mat)
  y <- x$vectors
  y <- y[, 1]
  z <- x$values[1]

  mod.mat <- matrix(z, k, k)
  mod.mat <- t(mod.mat * y) * y
  diag(mod.mat) <- 1
  mod.mat <- calc.cov.mat(mod.mat, sqrt(diag(cm)))
}

for(lvl in length(ind):1)
{
  fac.cov <- mod.mat
  mod.mat <- flm[[lvl]] %*% mod.mat %*% t(flm[[lvl]])
  sv <- diag(1 - mod.mat)
  diag(mod.mat) <- 1
  tr <- sqrt(diag(cov.mat[[lvl]]))
  mod.mat <- calc.cov.mat(mod.mat, tr)
}

if(!rm.sing.tkr)
{
  take <- colSums(ind[[1]]) == 1
  x <- diag(fac.cov)
  x[take] <- 0
  diag(fac.cov) <- x
  if(sum(take) > 1)
    take <- rowSums(ind[[1]][, take]) == 1
  else if (sum(take) == 1)
    take <- as.vector(ind[[1]][, take]) == 1
  else
    take <- rep(F, nrow(ind[[1]]))
  sv[take] <- 1
}

spec.risk <- tr * sqrt(sv)
fac.load <- tr * flm[[1]]

```

```

v <- flm[[1]] / sv
d <- solve(fac.cov) + t(flm[[1]]) %*% v
inv <- diag(1 / sv) - v %*% solve(d) %*% t(v)
inv <- calc.cov.mat(inv, 1 / tr)

result <- new.env()
result$spec.risk <- spec.risk
result$fac.load <- fac.load
result$fac.cov <- fac.cov
result$cov.mat <- mod.mat
result$inv.cov <- inv
result <- as.list(result)
return(result)
}

```

C R Code: Optimizer with Constraints & Bounds

In this appendix we give the R source code for the optimization algorithm with linear homogeneous constraints and position bounds we use in Section 6.5. This code is similar to the code for the bounded regression algorithm discussed in detail in (Kakushadze, 2015b) with one important difference, so our discussion here will be brief. The entry function is `bopt.calc.opt()`. The `args()` of `bopt.calc.opt()` are: `ret`, which is the N -vector of stock returns (for a given date); `load`, a matrix whose columns are the coefficients of the homogeneous constraints, so `dim(load)[1]` is N (e.g., if the sole constraint is the dollar neutrality constraint, then `load` is an $N \times 1$ matrix with unit elements); `inv.cov`, which is the $N \times N$ inverse factor model covariance matrix Γ_{ij}^{-1} ; `upper`, which is the N -vector of the upper bounds w_i^+ on the weights w_i (see below); `lower`, which is the N -vector of the lower bounds w_i^- on the weights w_i ; and `prec`, which is the desired precision with which the output weights w_i , the N -vector of which `bopt.calc.opt()` returns, must satisfy the normalization condition $\sum_{i=1}^N |w_i| = 1$. Here the weights are defined as $w_i \equiv H_i/I$ (the dollar holdings over the total investment level). See (Kakushadze, 2015b) for more detail.⁴⁴

```

bopt.calc.opt <- function (ret, load, inv.cov, upper, lower, prec = 1e-5)
{
  x <- bopt.gen.lm(ret, load, inv.cov)
  ret <- ret / sum(abs(x))
}

```

⁴⁴ The analog of the line `y <- t(load[Jt,]) %*% w.ret[Jt,]` below, reads `y <- t(w.load[Jt,]) %*% ret[Jt,]` in (Kakushadze, 2015b), where the analog of `inv.cov` is diagonal and both lines give the same result; however, for non-diagonal `inv.cov` here they do not.

```

repeat{
  x <- bopt.opt(ret, load, inv.cov, upper, lower)
  if(abs(sum(abs(x)) - 1) < prec)
    break

  ret <- ret / sum(abs(x))
}

return(x)
}

bopt.gen.lm <- function (x, y, z)
{
  if(is.vector(z))
    z <- diag(z)
  if(is.vector(y))
    y <- matrix(y, length(y), 1)
  if(is.vector(x))
    x <- matrix(x, length(x), 1)
  y1 <- z %*% y
  res <- (z - y1 %*% solve(t(y) %*% y1) %*% t(y1)) %*% x
  return(res)
}

bopt.opt <- function (ret, load, inv.cov, upper, lower, tol = 1e-6)
{
  calc.bounds <- function(z, x)
  {
    q <- x - z
    p <- rep(NA, length(x))
    pp <- pmin(x, upper)
    pm <- pmax(x, lower)
    p[q > 0] <- pp[q > 0]
    p[q < 0] <- pm[q < 0]
    t <- (p - z)/q
    t <- min(t, na.rm = T)
    z <- z + t * q
    return(z)
  }

  if(!is.matrix(load))
    load <- matrix(load, length(load), 1)

```

```

n <- nrow(load)
k <- ncol(load)

ret <- matrix(ret, n, 1)
upper <- matrix(upper, n, 1)
lower <- matrix(lower, n, 1)
z <- inv.cov
w.load <- z %*% load
w.ret <- z %*% ret

J <- rep(T, n)
Jp <- rep(F, n)
Jm <- rep(F, n)
z <- rep(0, n)

repeat{
  Jt <- J & !Jp & !Jm
  y <- t(load[Jt, ]) %*% w.ret[Jt, ]
  if(sum(Jp) > 1)
    y <- y + t(load[Jp, ]) %*% upper[Jp, ]
  else if(sum(Jp) == 1)
    y <- y + upper[Jp, ] * matrix(load[Jp, ], k, 1)
  if(sum(Jm) > 1)
    y <- y + t(load[Jm, ]) %*% lower[Jm, ]
  else if(sum(Jm) == 1)
    y <- y + lower[Jm, ] * matrix(load[Jm, ], k, 1)
  if(k > 1)
    take <- colSums(abs(load[Jt, ])) > 0
  else
    take <- T
  Q <- t(load[Jt, take]) %*% w.load[Jt, take]
  Q <- solve(Q)
  v <- Q %*% y[take]

  xJp <- Jp
  xJm <- Jm

  x <- w.ret - w.load[, take] %*% v
  x[Jp, ] <- upper[Jp, ]
  x[Jm, ] <- lower[Jm, ]

  z <- calc.bounds(z, x)
  Jp <- abs(z - upper) < tol

```

```

    Jm <- abs(z - lower) < tol

    if(all(Jp == xJp) & all(Jm == xJm))
      break
  }

  return(z)
}

```

D 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, Appendix B and Appendix C hereof is part of the copyrighted R code of Quantigic® Solutions LLC 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, Appendix B and Appendix C hereof and any and all copyrights therefor.

References

- Acharya, V.V. and Pedersen, L.H. (2005) Asset pricing with liquidity risk. *Journal of Financial Economics* 77(2): 375-410.
- Ang, A., Hodrick, R., Xing, Y. and Zhang, X. (2006) The Cross-Section of Volatility and Expected Returns. *Journal of Finance* 61(1): 259-299.
- Anson, M. (2013/14) Performance Measurement in Private Equity: Another Look at the Lagged Beta Effect. *The Journal of Private Equity* 17(1): 29-44.
- Asness, C.S. (1995) The Power of Past Stock Returns to Explain Future Stock Returns. Goldman Sachs Asset Management. Working paper.
- Asness, C. and Stevens, R. (1995) Intra- and Inter-Industry Variation in the Cross-Section of Expected Stock Returns. Goldman Sachs Asset Management. Working paper.
- Asness, C., Krail, R.J. and Liew, J.M. (2001) Do Hedge Funds Hedge? *The Journal of Portfolio Management* 28(1): 6-19.
- Bai, J. (2003) Inferential theory for factor models of large dimension. *Econometrica* 71(11): 135-171.
- Bai, J. and Li, K. (2012) Statistical Analysis of Factor Models of High Dimension. *The Annals of Statistics* 40(1): 436-465.
- Bai, J. and Ng, S. (2002) Determining the number of factors in approximate factor models. *Econometrica* 70(1): 191-221.
- Bansal, R. and Viswanathan, S. (1993) No Arbitrage and Arbitrage Pricing: A New Approach. *The Journal of Finance* 48(4): 1231-1262.
- Banz, R. (1981) The relationship between return and market value of common stocks. *Journal of Financial Economics* 9(1): 3-18.
- Basu, S. (1977) The investment performance of common stocks in relation to their price to earnings ratios: A test of the efficient market hypothesis. *Journal of Finance* 32(3): 663-682.
- Black, F. (1972) Capital market equilibrium with restricted borrowing. *Journal of Business* 45(3): 444-455.
- Black, F., Jensen, M. and Scholes, M. (1972) The capital asset pricing model: Some empirical tests. In: Jensen, M. (ed.) *Studies in the Theory of Capital Markets*. New York, NY: Praeger Publishers, pp. 79-121.

- Blume, O. and Friend, L. (1973) A new look at the capital asset pricing model. *Journal of Finance* 28(1): 19-33.
- Brandt, M.W., Brav, A., Graham, J.R. and Kumar, A. (2010) The idiosyncratic volatility puzzle: Time trend or speculative episodes? *Review of Financial Studies* 23(2): 863-899.
- Briner, B. and Connor, G. (2008) How much structure is best? A comparison of market model, factor model and unstructured equity covariance matrices. *Journal of Risk* 10(4), 3-30.
- Burmeister, E. and Wall, K.D. (1986) The arbitrage pricing theory and macroeconomic factor measures. *Financial Review* 21(1): 1-20.
- Campbell, J. (1987) Stock returns and the term structure. *Journal of Financial Economics* 18(2): 373-399.
- Campbell, J.Y., Lettau, M., Malkiel, B.G. and Xu, Y. (2001) Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. *Journal of Finance* 56(1): 1-43.
- Campbell, J. and Shiller, R. (1988) The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies* 1(3): 195-227.
- Carhart, M.M. (1997) Persistence in mutual fund performance. *Journal of Finance* 52(1): 57-82.
- Chamberlain, G. and Rothschild, M. (1983) Arbitrage, Factor Structure, and Mean-Variance Analysis on Large Asset Markets. *Econometrica* 51(5): 1281-1304.
- Chan, K.C., Chen, N. and Hsieh, D. (1985) An Exploratory Investigation of the Firm Size Effect. *Journal of Financial Economics* 14(3): 451-471.
- Chen, N., Grundy, B. and Stambaugh, R.F. (1990) Changing Risk, Changing Risk Premiums, and Dividend Yield Effects. *The Journal of Business* 63(1): 51-70.
- Chen, N., Roll, R. and Ross, S. (1986) Economic forces and the stock market. *Journal of Business* 59(3): 383-403.
- Chicheportiche, R. and Bouchaud, J.-P. (2014) A nested factor model for non-linear dependencies in stock returns. *Quantitative Finance* (forthcoming), DOI:10.1080/14697688.2014.994668.
- Cochrane, J.H. (2001) *Asset Pricing*. Princeton, NJ: Princeton University Press.

- Connor, G. (1984) A unified beta pricing theory. *Journal of Economic Theory* 34(1): 13-31.
- Connor, G. (1995) The Three Types of Factor Models: A Comparison of Their Explanatory Power. *Financial Analysts Journal* 51(3): 42-46.
- Connor, G. and Korajczyk, R. (1988) Risk and return in an equilibrium APT: Application of a new test methodology. *Journal of Financial Economics* 21(2): 255-289.
- Connor, G. and Korajczyk, R. (1989) An intertemporal beta pricing model. *Review of Financial Studies* 2(3): 373-392.
- Connor, G. and Korajczyk, R. (2010) Factor Models in Portfolio and Asset Pricing Theory. In: Guerard Jr, J.B. (ed.) *Handbook of Portfolio Construction: Contemporary Applications of Markowitz Techniques*. New York, NY: Springer, pp. 401-418.
- Daniel, K. and Titman, S. (1997) Evidence on the Characteristics of Cross Sectional Variation in Stock Returns. *Journal of Finance* 52(1): 1-33.
- DeBondt, W. and Thaler, R. (1985) Does the stock market overreact? *Journal of Finance* 40(3): 739-805.
- Dhrymes, P.J., Friend, I. and Gultekin, N.B. (1984) A Critical Reexamination of the Empirical Evidence on the Arbitrage Pricing Theory. *The Journal of Finance* 39(2): 323-346.
- Ederington, L. and Guan, W. (2002) Is implied volatility an informationally efficient and effective predictor of future volatility? *The Journal of Risk* 4(3): 29-46.
- Fama, E. and French, K. (1992) The cross-section of expected stock returns. *Journal of Finance* 47(2): 427-465.
- Fama, E.F. and French, K.R. (1993) Common risk factors in the returns on stocks and bonds. *J. Financ. Econ.* 33(1): 3-56.
- Fama, E. and French, K. (1996) Multifactor explanations for asset pricing anomalies. *Journal of Finance* 51(1): 55-94.
- Fama, E. and French, K. (2015) A Five-Factor Asset Pricing Model. *Journal of Financial Economics* (forthcoming), DOI: 10.1016/j.jfineco.2014.10.010.
- Fama, E.F. and MacBeth, J.D. (1973) Risk, Return and Equilibrium: Empirical Tests. *Journal of Political Economy* 81(3): 607-636.

- Ferson, W. and Harvey, C. (1991) The variation in economic risk premiums. *Journal of Political Economy* 99(2): 385-415.
- Ferson, W. and Harvey, C. (1999) Conditioning variables and the cross section of stock returns. *Journal of Finance* 54(4): 1325-1360.
- Forni, M., Hallin, M., Lippi, M. and Reichlin, L. (2000) The generalized dynamic factor model: Identification and estimation. *The Review of Economics and Statistics* 82(4): 540-554.
- Forni, M., Hallin, M., Lippi, M. and Reichlin, L. (2005) The generalized dynamic factor model: One-sided estimation and forecasting. *Journal of the American Statistical Association* 100(471): 830-840.
- Forni, M. and Lippi, M. (2001). The generalized dynamic factor model: Representation theory. *Econometric Theory* 17(6): 1113-1141.
- Goyal, A., Perignon, C. and Villa, C. (2008) How common are common return factors across the NYSE and Nasdaq? *Journal of Financial Economics* 90(3): 252-271.
- Goyal, A. and Santa-Clara, P. (2003) Idiosyncratic risk matters! *Journal of Finance* 58(3): 975-1007.
- Grinold, R.C. and Kahn, R.N. (2000) *Active Portfolio Management*. New York, NY: McGraw-Hill.
- Hall, A.D., Hwang, S. and Satchell, S.E. (2002) Using bayesian variable selection methods to choose style factors in global stock return models. *Journal of Banking and Finance* 26(12): 2301-2325.
- Haugen, R.A. (1995) *The New Finance: The Case Against Efficient Markets*. Upper Saddle River, NJ: Prentice Hall.
- Heaton, J. and Lucas, D.J. (1999) Stock Prices and Fundamentals. *NBER Macroeconomics Annual* 14(1): 213-242.
- Heston, S.L. and Rouwenhorst, K.G. (1994) Does Industrial Structure Explain the Benefits of International Diversification? *Journal of Financial Economics* 36(1): 3-27.
- Jagannathan, R. and Wang, Z. (1996) The conditional CAPM and the cross-section of expected returns. *Journal of Finance* 51(1): 3-53.
- Jegadeesh, N. and Titman, S. (1993) Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance* 48(1): 65-91.

- Jegadeesh, N. and Titman, S. (2001) Profitability of Momentum Strategies: An Evaluation of Alternative Explanations. *Journal of Finance* 56(2): 699-720.
- Kakushadze, Z. (2014a) 4-Factor Model for Overnight Returns. SSRN Working Papers Series, <http://ssrn.com/abstract=2511874> (October 19, 2014).
- Kakushadze, Z. (2014b) Russian-Doll Risk Models. SSRN Working Papers Series, <http://ssrn.com/abstract=2538123> (December 14, 2014).
- Kakushadze, Z. (2015a) Mean-Reversion and Optimization. *Journal of Asset Management* 16(1): 14-40.
- Kakushadze, Z. (2015b) Combining Alphas via Bounded Regression. SSRN Working Papers Series, <http://ssrn.com/abstract=2550335> (January 15, 2015).
- Kakushadze, Z. and Liew, J.K.-S. (2014) Custom v. Standardized Risk Models. SSRN Working Papers Series, <http://ssrn.com/abstract=2493379> (September 8, 2014).
- King, B.F. (1966) Market and Industry Factors in Stock Price Behavior. *Journal of Business* 39(1): 139-190.
- Korajczyk, R.A. and Sadka, R. (2008) Pricing the Commonality Across Alternative Measures of Liquidity. *Journal of Financial Economics* 87(1): 45-72.
- Kothari, S. and Shanken, J. (1997) Book-to-market, dividend yield and expected market returns: A time series analysis. *Journal of Financial Economics* 44(2): 169-203.
- Lakonishok, J., Shleifer, A. and Vishny, R.W. (1994) Contrarian Investment, Extrapolation, and Risk. *The Journal of Finance* 49(5): 1541-1578.
- Lee, J.-H. and Stefek, D. (2008) Do Risk Factors Eat Alphas? *The Journal of Portfolio Management* 34(4): 12-24.
- Lehmann, B. and Modest, D. (1988) The empirical foundations of the arbitrage pricing theory. *Journal of Financial Economics* 21(2): 213-254.
- Liew, J. and Vassalou, M. (2000) Can Book-to-Market, Size and Momentum be Risk Factors that Predict Economic Growth? *Journal of Financial Economics* 57(2): 221-245.
- Lintner, J. (1965) The valuation of risky assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics* 47(1): 13-37.
- Lo, A.W. (2010) *Hedge Funds: An Analytic Perspective*. Princeton, NJ: Princeton University Press.

- Lo, A.W. and MacKinlay, A.C. (1990) Data-snooping biases in tests of financial asset pricing models. *Review of Financial Studies* 3(3): 431-468.
- MacKinlay, A.C. (1995) Multifactor models do not explain deviations from the CAPM. *Journal of Financial Economics* 38(1): 3-28.
- MacQueen, J. (2003) The structure of multifactor equity risk models. *Journal of Asset Management* 3(4) 313-322.
- Markowitz, H.M. (1952) Portfolio Selection. *Journal of Finance* 7(1): 77-91.
- Markowitz, H.M. (1984) The Two-Beta Trap. *Journal of Portfolio Management* 11(1): 12-19.
- Menchero, J. and Mitra, I. (2008) The Structure of Hybrid Factor Models. *Journal of Investment Management* 6(3): 35-47.
- Merton, R. (1973) An intertemporal capital asset pricing model. *Econometrica* 41(5): 867-887.
- Miller, G. (2006) Needles, Haystacks, and Hidden Factors. *The Journal of Portfolio Management* 32(2): 25-32.
- Motta, G., Hafner, C. and von Sachs, R. (2011) Locally stationary factor models: identification and nonparametric estimation. *Econometric Theory* 27(6): 1279-1319.
- Mukherjee, D. and Mishra, A.K. (2005) Multifactor Capital Asset Pricing Model Under Alternative Distributional Specification. SSRN Working Papers Series, <http://ssrn.com/abstract=871398> (December 29, 2005).
- Ng, V., Engle, R.F. and Rothschild, M. (1992) A multi-dynamic-factor model for stock returns. *Journal of Econometrics* 52(1-2): 245-266.
- Pastor, L. and Stambaugh, R.F. (2003) Liquidity Risk and Expected Stock Returns. *The Journal of Political Economy* 111(3): 642-685.
- Roll, R. and Ross, S.A. (1980) An Empirical Investigation of the Arbitrage Pricing Theory. *Journal of Finance* 35(5): 1073-1103.
- Rosenberg, B. (1974) Extra-Market Components of Covariance in Security Returns. *Journal of Financial and Quantitative Analysis* 9(2): 263-274.
- Ross, S.A. (1976) The arbitrage theory of capital asset pricing. *Journal of Economic Theory* 13(3): 341-360.
- Ross, S.A. (1978a) A Simple Approach to the Valuation of Risky Streams. *Journal of Business* 51(3): 453-475.

- Ross, S.A. (1978b) Mutual Fund Separation in Financial Theory – The Separating Distributions. *Journal of Economic Theory* 17(2): 254-286.
- Scholes, M. and Williams, J. (1977) Estimating Betas from Nonsynchronous Data. *Journal of Financial Economics* 5(3): 309-327.
- Schwert, G. (1990) Stock returns and real activity: A century of evidence. *Journal of Finance* 45(4): 1237-1257.
- Shanken, J. (1987) Nonsynchronous data and the covariance-factor structure of returns. *Journal of Finance* 42(2): 221-231.
- Shanken, J. (1990) Intertemporal Asset Pricing: An Empirical Investigation. *Journal of Econometrics* 45(1-2): 99-120.
- Shanken, J. and Weinstein, M.I. (2006) Economic Forces and the Stock Market Revisited. *Journal of Empirical Finance* 13(2): 129-144.
- Sharpe, W.F. (1963) A simplified model for portfolio analysis. *Management Science* 9(2): 277-293.
- Sharpe, W. (1964) Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance* 19(3): 425-442.
- Stock, J.H. and Watson, M.W. (2002a) Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics* 20(2): 147-162.
- Stock, J.H. and Watson, M.W. (2002b) Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association* 97(460): 1167-1179.
- Stroyny, A.L. (2005) Estimating a Combined Linear Factor Model. In: Knight, J. and Satchell, S.E. (eds.) *Linear Factor Models in Finance*. Oxford: Elsevier Butterworth-Heinemann.
- Treynor, J.L. (1999) Towards a Theory of Market Value of Risky Assets. In: Korajczyk, R. (ed.) *Asset Pricing and Portfolio Performance: Models, Strategy, and Performance Metrics*. London: Risk Publications.
- Vassalou, M. (2003) News Related to Future GDP Growth as a Risk Factor in Equity Returns. *Journal of Financial Economics*. 68(1): 47-73.
- Whitelaw, R. (1997) Time variations and covariations in the expectation and volatility of stock market returns. *Journal of Finance* 49(2): 515-541.
- Zangari, P. (2003) Equity factor risk models. In: Litterman, B. (ed.) *Modern Investment Management: An Equilibrium Approach*. New York, NY: John Wiley & Sons, Inc., pp. 334-395.

Zhang, C. (2010) A Re-examination of the Causes of Time-varying Stock Return Volatilities. *Journal of Financial and Quantitative Analysis* 45(3): 663-684.

Table 1: First column: the number of principal components K ; last column: $g(K)$ defined in (30); Min, 1st Quartile, Median, Mean, 3rd Quartile and Max refer to the corresponding quantities for the ratio $\tilde{\xi}_i^2 = \xi_i^2/C_{ii}$ (specific variance over total variance). The number of observations (days) in the time series is $M + 1 = 20$. The number of (randomly selected) stocks is $N = 2316$. All quantities are rounded to 3 digits. The value of K fixed via (29) is $K = 12$. See Figure 1 for a density plot.

K	Min	1st Quartile	Median	Mean	3rd Quartile	Max	$g(K)$
1	0.16	0.62	0.824	0.771	0.953	1	0.4
2	0.137	0.525	0.693	0.682	0.867	1	0.37
3	0.084	0.453	0.629	0.618	0.8	0.999	0.29
4	0.075	0.405	0.562	0.56	0.718	0.992	0.27
5	0.06	0.355	0.501	0.51	0.668	0.981	0.235
6	0.06	0.312	0.449	0.462	0.606	0.977	0.233
7	0.057	0.272	0.396	0.417	0.552	0.931	0.203
8	0.033	0.233	0.347	0.375	0.503	0.916	0.139
9	0.029	0.203	0.306	0.334	0.446	0.884	0.111
10	0.019	0.172	0.264	0.294	0.39	0.84	0.056
11	0.01	0.144	0.227	0.256	0.339	0.84	0.018
12	0.009	0.118	0.194	0.22	0.294	0.84	0.013
13	0.008	0.093	0.157	0.186	0.25	0.728	0.056
14	0.002	0.07	0.122	0.152	0.204	0.696	0.12
15	0.002	0.05	0.089	0.119	0.162	0.686	0.128
16	0	0.029	0.062	0.088	0.122	0.608	0.211
17	0	0.014	0.035	0.057	0.077	0.606	0.221
18	0	0.003	0.011	0.028	0.034	0.592	0.231

Table 2: Simulation results for the weighted regression alphas discussed in Section 6.4 and the optimized alphas discussed in Section 6.5, without any bounds on the dollar holdings. All quantities are rounded to 2 digits. See Figure 2 for P&L plots.

Alpha	ROC	SR	CPS
Regression: Principal Components	46.80%	11.50	2.05
Optimization: Principal Components	47.74%	11.88	2.26
Regression: BICS Sub-Industries	49.36%	12.89	2.16
Regression: Heterotic Risk Factors	51.89%	13.63	2.27
Optimization: Heterotic Risk Model	55.90%	15.41	2.67

Table 3: Simulation results for the weighted regression alphas discussed in Section 6.4 and the optimized alphas discussed in Section 6.5, with the liquidity bounds (89) on the dollar holdings. All quantities are rounded to 2 digits. See Figure 3 for P&L plots.

Alpha	ROC	SR	CPS
Regression: Principal Components	41.27%	14.24	1.84
Optimization: Principal Components	40.92%	14.33	1.96
Regression: BICS Sub-Industries	44.56%	16.51	1.97
Regression: Heterotic Risk Factors	46.86%	18.30	2.08
Optimization: Heterotic Risk Model	49.00%	19.23	2.36

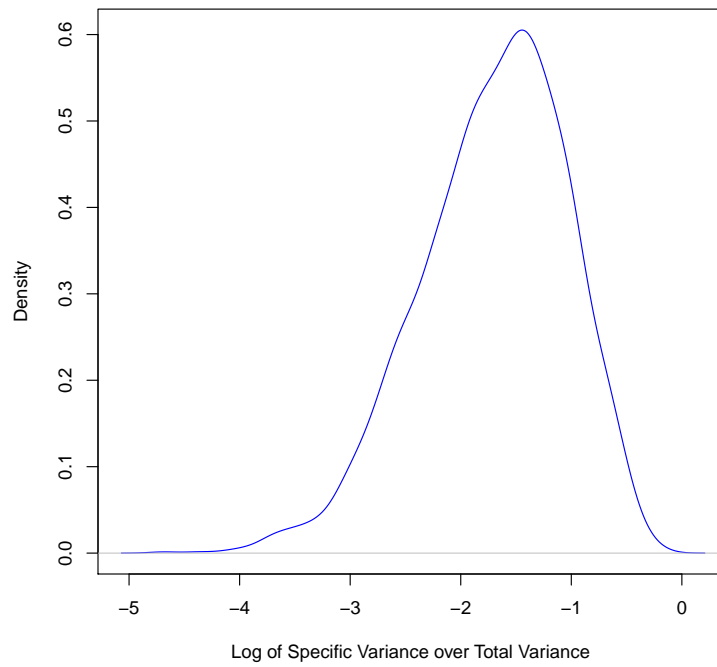


Figure 1. The density (computed using the R function `density()`) for the log of the ratio $\tilde{\xi}_i^2 = \xi_i^2 / C_{ii}$ (specific variance over total variance) for the $K = 12$ case in Table 1.

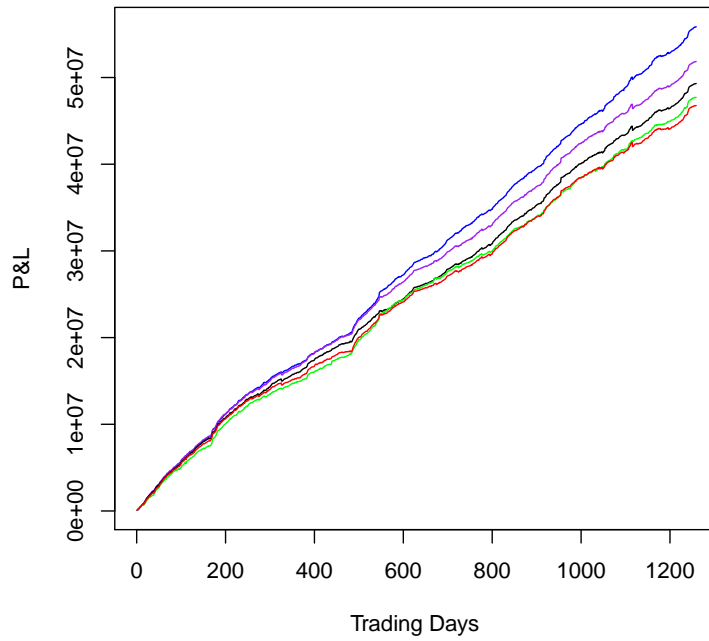


Figure 2. P&L graphs for the intraday alphas without liquidity bounds summarized in Table 2. Bottom-to-top-performing: i) Regression over 20 principal components (Section 6.4), ii) optimization using the principal component risk model (Section 6.5), iii) regression over the BICS sub-industries (Section 6.4), iv) regression over the heterotic risk factor loadings (Section 6.4), and v) optimization using the heterotic risk model (Section 6.5). The investment level is \$10M long plus \$10M short.

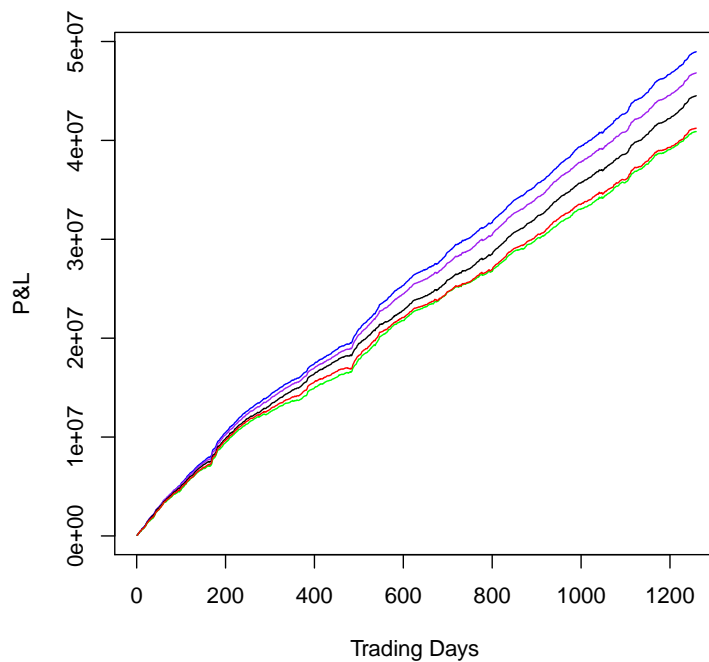


Figure 3. P&L graphs for the intraday alphas with liquidity bounds summarized in Table 3. Bottom-to-top-performing: i) optimization using the principal component risk model (Section 6.5), ii) Regression over 20 principal components (Section 6.4), iii) regression over the BICS sub-industries (Section 6.4), iv) regression over the heterotic risk factor loadings (Section 6.4), and v) optimization using the heterotic risk model (Section 6.5). The investment level is \$10M long plus \$10M short.