



NUS RMI Industrial Research Papers – No. 2022-01

The Scheweinler-Wigner Orthogonalization For Risk Model Construction

Chang CAO and Kan CHEN

June 2022

NUS Risk Management Institute

21 HENG MUI KENG TERRACE, #04-03 I3 BUILDING, SINGAPORE 119613

www.rmi.nus.edu.sg/research/industrial-research-papers

THE SCHEWEINLER-WIGNER ORTHOGONALIZATION FOR RISK MODEL CONSTRUCTION

Chang Cao and Kan Chen

Risk Management Institute

National University of Singapore

Abstract

In this paper, we investigate the Schweinler–Wigner/Löwdin (SW/L) orthogonalization scheme for risk model constructions. The SW/L scheme maximizes the overlaps between the original risk factors and the new orthogonalized (independent) risk factors. We test the robustness of the risk factor orthogonalization procedure on four risk models based on combinations of Farmer-French 3 factors and various sets of industry factors in both the US and the Chinese markets. We show that the original risk factors and the corresponding new orthogonalized factors are highly correlated, and such high correlations are stable over a long period. This allows the orthogonalized factors to retain the economic intuition of original explicitly-defined risk factors, leading to a better and clearer risk attribution.

1. Introduction

1.1 Systematic risk

Systematic risk, also known as “un-diversifiable risk”, refers to the risk inherent to the entire market or market segment. It affects the overall market, not just a particular asset. This type of risk is both unpredictable and impossible to avoid. Systemic risks include policy risks, macroeconomic risks, and environmental risks. Systematic risk cannot be mitigated through diversification; hedging is thus the main tool for mitigating systematic risk. To hedge the risk effectively requires a risk model for identifying risk factors and performing quantitative risk attribution.

1.2 The necessity of orthogonalization

In asset pricing and portfolio management, a multi-factor risk model is often used to explain stock returns in terms of various systematic risk factors. Under the multi-factor model framework, the systematic risk in the return of a risky asset is explained as a linear combination of the risk factors,

$$r_i(t) = \sum_k \beta_i^k f_k(t) + \epsilon_i(t),$$

where $r_i(t)$ is the return (from t-1 to t) of the asset i, $f_k(t)$ is the risk factor portfolio return, β_i^k is factor loading, and $\epsilon_i(t)$ represents the idiosyncratic part of the asset return.

If the factors are linearly independent, the factor loading determines how much systematic risk is loaded on the risk factor. Yet, the collinearity between risk factors is not normally eliminated in a risk model. This is the case for many well-known risk models, such as Barra’s risk model, which is constructed with risk factors explicitly defined. There are also risk models based on the principal component analysis (PCA); in these models, the risk factors are the principal components, and they are linearly independent. The PCA-based model, however, has a disadvantage in that the risk factors don’t have clearly defined meaning; it is difficult to perform proper hedging of such risk factors. In this paper, we study a risk model construction method using the risk factors that are both linearly independent and interpretable. We start with a set of well-defined risk factors, then employ an orthogonalization scheme to generate a new set of linearly independent factors. The crucial choice here is to use an orthogonalization method that maximizes the overlap between the new factors and the original factors we start with; a significant overlap would allow the new factors to retain the economic intuition of the original factors. Schweinler–Wigner/Löwdin (SW/L) orthogonalization method (Schweinler-Wigner, 1970; Löwdin 1950) is precisely the one we need for this purpose.

2. Orthogonalization procedure

2.1 Literature review

There have been many studies of risk factors for analyzing stock performance. The Fama-French 3-factor model (FF3, 1993) includes three well-known factors: the

market, size, and value factors. Jegadeesh & Titman (1993) published the first research on the momentum factor. Chen and De Bondt(2004) further found that investors could benefit from chasing investment styles. In addition to including style factors, a good risk model also need to include factors related to market and industry exposures. In this paper, we study risk models that include both the FF3 risk factors and the factors related to sector/industry exposures.

Under a multi-factor framework, the possible covariance between the risk factors makes it complicated to attribute the sources of risk (Campbell and Mei, 1993). To eliminate the covariances among the risk factors, we need to use an orthogonalization method. The Gram- Schmidt orthogonalization method is perhaps the most used, but the Schweinler–Wigner/Löwdin method is becoming recognized in finance research. Klein and Chow (2013) use the SW/L method for factor orthogonalization in the context of examining the contribution of the Fama and French (1993) three factors to the return of risky assets. Klein(2011) found that factors other than market and size factors play an important role in explaining portfolios' volatility. Adcock et al. (2019) found that a small number of variables, including the change in the term premium, unexpected inflation, and change in volatility dominate non-market explained variation.

2.2 Orthogonalization methods

We used two orthogonalization methods, namely, the Gram-Schmidt (GS) method and Schweinler–Wigner/Löwdin (SW/L) method, and compared their performance when the number of risk factors increases.

2.2.1 Gram-Schmidt process

The Gram–Schmidt process is a method for ortho-normalizing a set of vectors in an inner product space, most commonly the Euclidean space R_n equipped with the standard inner product. The process works as follows, starting with a set of vectors $\{v_i, i = 1, \dots, k\}$:

$$\begin{aligned} u_1 &= v_1 \\ u_2 &= v_2 - \text{proj}_{u_1}(v_2) \\ u_3 &= v_3 - \text{proj}_{u_1}(v_3) - \text{proj}_{u_2}(v_3) \\ &\dots \\ u_k &= v_k - \sum_{j=1}^{k-1} \text{proj}_{u_j}(v_k) \end{aligned}$$

Here $\text{proj}_u(v)$ denotes the projection of the vector v to u .

The vectors u_1, u_2, \dots, u_k are the orthogonal set of vectors generated in this process.

2.2.2 Schweinler–Wigner/Löwdin (1970) symmetric procedure

The Schweinler–Wigner procedure is a method for orthogonalizing a set of vectors (arranged as the columns of a matrix in this procedure) symmetrically. It has the advantage that it is not sensitive to the order in which vectors are arranged. The procedure maximizes the overlap between the original vectors and the

orthogonalized vectors generated. The procedure treats all factors on an equal footing. The process works as follows:

F_{TK} denotes the original factor values from factor 1 to k and time 1 to T .

We want to find a matrix S_{KK} such that

$$G_{TK} = F_{TK}S_{KK},$$

where G_{TK} is the required set of orthogonal vectors. The orthonormal condition of G_{TK} means that $G^T G = I$, $S^T F^T F S = I$, or $S^T M S = I$, where I is the identity matrix and M is the Gram matrix of the original set.

By diagonalizing M , we can find O_{KK} , D_{KK} , such that

$$M_{KK} = O_{KK}D_{KK}O_{KK}^T,$$

It is easy to see that $S^T M S = I$ can be solved by constructing

$$S_{KK} = O_{KK}D_{KK}^{-1/2}O_{KK}^T$$

3. The US market

3.1 The 14-factor model

In this session, we consider a 14-factor risk model in the US stock market, which include the well-known Fama-French three factors, together with eleven sector factors based on Global Industry Classification Standard (GICS)– Energy, Materials, Industrials, Consumer Discretionary, Consumer Staples, Health Care, Financials, Information Technology, Communication Services, Utilities, and Real Estate. The universe we consider consists of the component stocks of the SP500 index. We use the monthly returns of stocks from January 2003 to May 2021 to construct the risk model. Note that the market component is removed from the sector returns, thus the correlation between the market risk factor and the sector risk factor is zero.

3.1.1 Orthogonalization performance

The tables below show the correlation between factors in the same set. We can see that the correlations are quite significant among the original factors; the orthogonalization eliminates the correlations among the orthogonalized factors as expected.

	Mkt-RF	SMB	HML	Energy	Materials	Industrials	Cons Disc	Cons Stap	Healthcare	Financials	IT	Comm	Utilities	Real Estate
Mkt-RF	1.00000	0.37367	0.26729	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
SMB	0.37367	1.00000	0.15744	0.10674	0.01603	-0.00670	0.15822	-0.43301	-0.14764	-0.09183	-0.03617	-0.21774	-0.21698	0.04008
HML	0.26729	0.15744	1.00000	0.22876	0.06446	0.28801	-0.14439	-0.05353	-0.21693	0.64774	-0.41968	-0.18264	-0.05173	0.16311
Energy	0.00000	0.10674	0.22876	1.00000	0.25446	-0.02856	-0.32608	-0.19353	-0.20117	-0.11960	-0.23034	-0.04288	0.00085	-0.17960
Materials	0.00000	0.01603	0.06446	0.25446	1.00000	0.28534	-0.03809	-0.17291	-0.18603	-0.04312	-0.03461	-0.07659	-0.11710	0.06241
Industrials	0.00000	-0.00670	0.28801	-0.02856	0.28534	1.00000	-0.01455	0.05419	-0.15669	0.24492	-0.20616	-0.11566	-0.09148	0.10085
Cons Disc	0.00000	0.15822	-0.14439	-0.32608	-0.03809	-0.01455	1.00000	-0.02854	-0.13787	-0.07257	0.18587	0.09596	-0.22896	0.19483
Cons Stap	0.00000	-0.43301	-0.05353	-0.19353	-0.17291	0.05419	-0.02854	1.00000	0.26097	-0.06079	-0.16230	0.14591	0.41183	0.13199
Healthcare	0.00000	-0.14764	-0.21693	-0.20117	-0.18603	-0.15669	-0.13787	0.26097	1.00000	-0.08644	-0.25697	-0.12670	0.12821	0.03628
Financials	0.00000	-0.09183	0.64774	-0.11960	-0.04312	0.24492	-0.07257	-0.06079	-0.08644	1.00000	-0.29951	-0.22419	-0.25551	0.10201
IT	0.00000	-0.03617	-0.41968	-0.23034	-0.03461	-0.20616	0.18587	-0.16230	-0.25697	-0.29951	1.00000	0.09079	-0.12590	-0.07676
Comm	0.00000	-0.21774	-0.18264	-0.04288	-0.07659	-0.11566	0.09596	0.14591	-0.12670	-0.22419	0.09079	1.00000	0.10573	-0.06461
Utilities	0.00000	-0.21698	-0.05173	0.00085	-0.11710	-0.09148	-0.22896	0.41183	0.12821	-0.25551	-0.12590	0.10573	1.00000	0.30151
Real Estate	0.00000	0.04008	0.16311	-0.17960	0.06241	0.10085	0.19483	0.13199	0.03628	0.10201	-0.07676	-0.06461	0.30151	1.00000

Table 1 Correlation of original factors

	Mkt-RF	SMB	HML	Energy	Materials	Industrials	Cons Disc	Cons Stap	Healthcare	Financials	IT	Comm	Utilities	Real Estate
Mkt-RF	1.00000	-0.00050	0.02084	-0.00487	-0.00016	-0.00737	0.00193	0.00032	0.00384	-0.02332	0.00840	0.00110	-0.00608	-0.00512
SMB	-0.00050	1.00000	0.00022	-0.00005	0.00000	-0.00008	0.00002	0.00000	0.00004	-0.00025	0.00009	0.00001	-0.00007	-0.00005
HML	0.02084	0.00022	1.00000	0.00218	0.00007	0.00330	-0.00086	-0.00014	-0.00172	0.01044	-0.00376	-0.00049	0.00272	0.00229
Energy	-0.00487	-0.00005	0.00218	1.00000	-0.00002	-0.00077	0.00020	0.00003	0.00040	-0.00244	0.00088	0.00011	-0.00063	-0.00053
Materials	-0.00016	0.00000	0.00007	-0.00002	1.00000	-0.00003	0.00001	0.00000	0.00001	-0.00008	0.00003	0.00000	-0.00002	-0.00002
Industrials	-0.00737	-0.00008	0.00330	-0.00077	-0.00003	1.00000	0.00030	0.00005	0.00061	-0.00369	0.00133	0.00017	-0.00096	-0.00081
Cons Disc	0.00193	0.00002	-0.00086	0.00020	0.00001	0.00030	1.00000	-0.00001	-0.00016	0.00096	-0.00035	-0.00005	0.00025	0.00021
Cons Stap	0.00032	0.00000	-0.00014	0.00003	0.00000	0.00005	-0.00001	1.00000	-0.00003	0.00016	-0.00006	-0.00001	0.00004	0.00003
Healthcare	0.00384	0.00004	-0.00172	0.00040	0.00001	0.00061	-0.00016	-0.00003	1.00000	0.00192	-0.00069	-0.00009	0.00050	0.00042
Financials	-0.02332	-0.00025	0.01044	-0.00244	-0.00008	-0.00369	0.00096	0.00016	0.00192	1.00000	0.00420	0.00055	-0.00304	-0.00256
IT	0.00840	0.00009	-0.00376	0.00088	0.00003	0.00133	-0.00035	-0.00006	-0.00069	0.00420	1.00000	-0.00020	0.00110	0.00092
Comm	0.00110	0.00001	-0.00049	0.00011	0.00000	0.00017	-0.00005	-0.00001	-0.00009	0.00055	-0.00020	1.00000	0.00014	0.00012
Utilities	-0.00608	-0.00007	0.00272	-0.00063	-0.00002	-0.00096	0.00025	0.00004	0.00050	-0.00304	0.00110	0.00014	1.00000	-0.00067
Real Estate	-0.00512	-0.00005	0.00229	-0.00053	-0.00002	-0.00081	0.00021	0.00003	0.00042	-0.00256	0.00092	0.00012	-0.00067	1.00000

Table 2 Correlation of orthogonalized factors using the GS method

	Mkt-RF	SMB	HML	Energy	Materials	Industrials	Cons Disc	Cons Stap	Healthcare	Financials	IT	Comm	Utilities	Real Estate
Mkt-RF	1.00000	-0.00782	0.01702	-0.00761	0.00017	-0.00574	0.00267	-0.00106	0.00705	-0.02311	0.00967	0.00211	-0.00104	-0.00481
SMB	-0.00782	1.00000	0.00270	-0.00121	0.00003	-0.00091	0.00042	-0.00017	0.00112	-0.00367	0.00153	0.00034	-0.00016	-0.00076
HML	0.01702	0.00270	1.00000	0.00263	-0.00006	-0.00198	-0.00092	0.00037	-0.00243	0.00798	-0.00334	-0.00073	0.00036	0.00166
Energy	-0.00761	-0.00121	0.00263	1.00000	0.00003	-0.00089	0.00041	-0.00016	0.00109	-0.00357	0.00149	0.00033	-0.00016	-0.00074
Materials	0.00017	0.00003	-0.00006	0.00003	1.00000	0.00002	-0.00001	0.00000	-0.00002	0.00008	-0.00003	-0.00001	0.00000	0.00002
Industrials	-0.00574	-0.00091	0.00198	-0.00089	0.00002	1.00000	0.00031	-0.00012	0.00082	-0.00269	0.00113	0.00025	-0.00012	-0.00056
Cons Disc	0.00267	0.00042	-0.00092	0.00041	-0.00001	0.00031	1.00000	0.00006	-0.00038	0.00125	-0.00052	-0.00011	0.00006	0.00026
Cons Stap	-0.00106	-0.00017	0.00037	-0.00016	0.00000	-0.00012	0.00006	1.00000	0.00015	-0.00050	0.00021	0.00005	-0.00002	-0.00010
Healthcare	0.00705	0.00112	-0.00243	0.00109	-0.00002	0.00082	-0.00038	0.00015	1.00000	0.00330	-0.00138	-0.00030	0.00015	0.00069
Financials	-0.02311	-0.00367	0.00798	-0.00357	0.00008	-0.00269	0.00125	-0.00050	0.00330	1.00000	0.00453	0.00099	-0.00049	-0.00225
IT	0.00967	0.00153	-0.00334	0.00149	-0.00003	0.00113	-0.00052	0.00021	-0.00138	0.00453	1.00000	-0.00041	0.00020	0.00094
Comm	0.00211	0.00034	-0.00073	0.00033	-0.00001	0.00025	-0.00011	0.00005	-0.00030	0.00099	-0.00041	1.00000	0.00004	0.00021
Utilities	-0.00104	-0.00016	0.00036	-0.00016	0.00000	-0.00012	0.00006	-0.00002	0.00015	-0.00049	0.00020	0.00004	1.00000	-0.00010
Real Estate	-0.00481	-0.00076	0.00166	-0.00074	0.00002	-0.00056	0.00026	-0.00010	0.00069	-0.00225	0.00094	0.00021	-0.00010	1.00000

Table 3 Correlation of orthogonalized factors using the SW/L method

What is important for our purpose is whether the orthogonalized factors are closely related to the original as we aim to retain the most economic intuition of the original factors. The table below shows the correlation between original and orthogonalized factors using the GS and SW/L schemes.

	Mkt-RF	SMB	HML	Energy	Materials	Industrials	Cons Disc	Cons Stap	Healthcare	Financials	IT	Comm	Utilities	Real Estate
GS	1.00000	0.92738	0.96699	0.96676	0.96698	0.90096	0.91676	0.85755	0.90186	0.66067	0.76594	0.89659	0.81304	0.84494
SW/L	0.98524	0.96791	0.98749	0.94171	0.96822	0.92781	0.92397	0.83846	0.91628	0.68904	0.86180	0.94235	0.92578	0.95786

Table 4 Correlation between original and orthogonalized factors

We can see that the SW/L method outperforms the GS method in this measure. Using the GS method, six correlation coefficients are below 0.9 whereas using the SW/L method, three correlation coefficients are below 0.9, with the lowest value being 0.69 for the Financials factor. The low value of the correlation for the financials factor is due to the high correlation of the financials factor to other factors in the original set (for example, the financials factor has a high correlation value of 0.65 with HML). But overall the SW/L method is a better orthogonalization scheme for generating a bigger overlap of the original and the orthogonalized factors.

3.1.2 Comparing the SW/L method over multiple periods

For the SW/L method, we would like to study the stability of the coefficient matrix relating old factors to new factors over multiple periods. This is important for a robust risk model construction using the orthogonalization scheme. Over the two 5-year periods (2011-2015, 2016-2020), we check that the coefficient matrix is indeed

quite stable. The correlation coefficient between $S_{2011-2015}$ and $S_{2016-2020}$ is 0.9249. The orthogonalized factors generated from the SW/L method are thus stable over the two five-year periods.

Regressing the performance of individual stocks on the fourteen factors, the R^2 is 0.4856 on average. The systematic risk accounts for 48.56% of the total variance in stock performance.

3.2 The 27-factor model

To construct the 27-factor model, we chose the Fama-French three factors, together with twenty-four industry group factors based on Global Industry Classification Standard (GICS). we used the GICS code to represent the industry group as shown in Table 5. Note that the market component is removed from the industry group returns

	Industry Group
1010	Energy
1510	Materials
2010	Capital Goods
2020	Commercial & Professional Services
2030	Transportation
2510	Automobiles & Components
2520	Consumer Durables & Apparel
2530	Consumer Services
2550	Retailing
3010	Food & Staples Retailing
3020	Food, Beverage & Tobacco
3030	Household & Personal Products
3510	Health Care Equipment & Services
3520	Pharmaceuticals, Biotechnology & Life Sciences
4010	Banks
4020	Diversified Financials
4030	Insurance
4510	Software & Services
4520	Technology Hardware & Equipment
4530	Semiconductors & Semiconductor Equipment
5010	Communication Services
5020	Media & Entertainment
5510	Utilities
6010	Real Estate

Table 5 GICS industry group code

3.2.1 Orthogonalization performance

The tables below show the correlation between factors in the same set. We can see that the correlation is mostly eliminated after orthogonalization using both methods.

Mkt-RF	SMB	HML	1010	1510	2010	2020	2030	2510	2520	2530	2550	3010	3020	3030	3510	3520	4010	4020	4030	4510	4520	4530	5010	5020	5510	6010		
Mkt-RF	1.00000	0.26279	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	
SMB	0.26279	1.00000	0.12744	0.10614	0.01803	-0.01574	0.02482	0.01717	0.08991	0.18553	0.03801	0.19996	0.03266	-0.07103	-0.31862	-0.15446	0.09219	-0.11972	-0.18358	0.02119	0.15221	-0.20261	-0.24048	-0.21688	0.14908	0.00000	0.00000	0.00000
HML	0.00000	0.00000	0.22827	0.06446	0.22827	0.06446	0.22827	0.06446	0.22827	0.06446	0.22827	0.06446	0.22827	0.06446	0.22827	0.06446	0.22827	0.06446	0.22827	0.06446	0.22827	0.06446	0.22827	0.06446	0.22827	0.06446	0.22827	0.06446

Table 6 Correlation of original factors

Mkt-RF	SMB	HML	1010	1510	2010	2020	2030	2510	2520	2530	2550	3010	3020	3030	3510	3520	4010	4020	4030	4510	4520	4530	5010	5020	5510	6010		
Mkt-RF	1.00000	0.26279	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	
SMB	0.26279	1.00000	0.12744	0.10614	0.01803	-0.01574	0.02482	0.01717	0.08991	0.18553	0.03801	0.19996	0.03266	-0.07103	-0.31862	-0.15446	0.09219	-0.11972	-0.18358	0.02119	0.15221	-0.20261	-0.24048	-0.21688	0.14908	0.00000	0.00000	0.00000
HML	0.00000	0.00000	0.22827	0.06446	0.22827	0.06446	0.22827	0.06446	0.22827	0.06446	0.22827	0.06446	0.22827	0.06446	0.22827	0.06446	0.22827	0.06446	0.22827	0.06446	0.22827	0.06446	0.22827	0.06446	0.22827	0.06446	0.22827	0.06446

Table 7 Correlation of orthogonalized factors using the GS method

Mkt-RF	SMB	HML	1010	1510	2010	2020	2030	2510	2520	2530	2550	3010	3020	3030	3510	3520	4010	4020	4030	4510	4520	4530	5010	5020	5510	6010	
Mkt-RF	1.00000	0.26279	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	
SMB	0.26279	1.00000	0.12744	0.10614	0.01803	-0.01574	0.02482	0.01717	0.08991	0.18553	0.03801	0.19996	0.03266	-0.07103	-0.31862	-0.15446	0.09219	-0.11972	-0.18358	0.02119	0.15221	-0.20261	-0.24048	-0.21688	0.14908	0.00000	0.00000
HML	0.00000	0.00000	0.22827	0.06446	0.22827	0.06446	0.22827	0.06446	0.22827	0.06446	0.22827	0.06446	0.22827	0.06446	0.22827	0.06446	0.22827	0.06446	0.22827	0.06446	0.22827	0.06446	0.22827	0.06446	0.22827	0.06446	0.22827

Table 8 Correlation of orthogonalized factors using the SW/L method

	Mkt-RF	SMB	HML	1010	1510	2010	2020	2030	2510	2520	2530	2550	3010	3020	3030	3510	3520	4010	4020	4030	4510	4520	4530	5010	5020	5510	6010
GS	1.00000	0.92738	0.96699	0.96676	0.96698	0.90958	0.98369	0.89279	0.95957	0.89999	0.92202	0.83493	0.86390	0.85365													
SW/L	0.98524	0.96786	0.98743	0.91235	0.93689	0.89598	0.94533	0.93319	0.98328	0.89028	0.94917	0.90159	0.92406	0.83619													
		3030	3510	3520	4010	4020	4030	4510	4520	4530	5010	5020	5510	6010													
GS	0.77605	0.91538	0.79780	0.61474	0.77267	0.77686	0.81232	0.79032	0.83513	0.84665	0.82152	0.75518	0.80469														
SW/L	0.87801	0.90657	0.87792	0.67107	0.83082	0.85872	0.85947	0.89215	0.92875	0.93819	0.93025	0.88847	0.94178														

Table 9 Correlation between original and orthogonalized factors

Again the SW/L method outperforms the GS method in this overlap measure. Using the GS method, out of the twenty-seven correlations between new and original industry groups, seven correlations are below 0.80 with the lowest value being 0.61.

Using the SW/L method, twenty-six out of twenty-seven-factor correlations are above 0.80, with the only exception being 0.67 for the Banks industry group (4010). These are likely due to the relatively high correlation between Banks and other factors. Overall the outperformance of the SW/L is more pronounced using this larger set of factors.

3.2.2 Comparing the SW/L method over multiple periods

Using the SW/L method, we further investigate whether the coefficient matrix relating old factors to new factors is similar over multiple periods.

Over the two 5-year periods (2011-2015, 2016-2020), the coefficient matrix is stable. The correlation coefficient between $S_{2011-2015}$ and $S_{2016-2020}$ is 0.9498. On average, 57.29% of the total variance can be explained by the model, which is an improvement compared to the 14-factor model.

4. The China market

4.1 The 14-factor model

We perform a similar study in the context of the Chinese market, using the monthly returns of stocks from January 2006 to May 2021 May. The universe used consists of the component stocks for CSI 300. As for the US market, the fourteen factors are the Fama-French three factors plus eleven sector factors based on Global Industry Classification Standard (GICS).

4.1.1 Orthogonalization performance

The tables below show the correlation between factors in the same set.

	Mkt-RF	SMB	HML	Energy	Materials	Industrials	Cons Disc	Cons Stap	Healthcare	Financials	IT	Comm	Utilities	Real Estate
Mkt-RF	1.00000	0.10884	-0.17641	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
SMB	0.10884	1.00000	-0.26578	-0.29799	0.25394	0.30683	0.43739	0.06464	0.42801	-0.53977	0.52925	0.18550	0.24169	0.04924
HML	-0.17641	-0.26578	1.00000	0.41847	-0.08912	-0.08892	-0.37708	-0.48999	-0.59235	0.47054	-0.52723	-0.42738	0.08301	0.22374
Energy	0.00000	-0.29799	0.41847	1.00000	0.12149	0.07853	-0.26937	-0.22586	-0.39521	0.36669	-0.28916	-0.20942	0.11513	-0.00057
Materials	0.00000	0.25394	-0.08912	0.12149	1.00000	0.18763	0.12442	-0.04232	0.01964	-0.22344	0.18188	0.07453	0.14172	-0.10054
Industrials	0.00000	0.30683	-0.08892	0.07853	0.18763	1.00000	0.19775	0.03258	0.15058	-0.21622	0.30505	0.29701	0.41480	-0.08378
Cons Disc	0.00000	0.43739	-0.37708	-0.26937	0.12442	0.19775	1.00000	0.32911	0.47690	-0.58398	0.42552	0.22123	0.13102	-0.09229
Cons Stap	0.00000	0.06464	-0.48999	-0.22586	-0.04232	0.03258	0.32911	1.00000	0.48563	-0.21044	0.20268	0.15440	-0.04209	-0.17808
Healthcare	0.00000	0.42801	-0.59235	-0.39521	0.01964	0.15058	0.47690	0.48563	1.00000	-0.51835	0.59415	0.30578	0.15877	-0.20969
Financials	0.00000	-0.53977	0.47054	0.36669	-0.22344	-0.21622	-0.58398	-0.21044	-0.51835	1.00000	0.46547	-0.12714	-0.15521	0.11344
IT	0.00000	0.52925	-0.52723	-0.28916	0.18188	0.30505	0.42552	0.20268	0.59415	-0.46547	1.00000	0.53067	0.02921	-0.15282
Comm	0.00000	0.18550	-0.42738	-0.20942	0.07453	0.29701	0.22123	0.15440	0.30578	-0.12714	0.53067	1.00000	0.10360	-0.09055
Utilities	0.00000	0.24169	0.08301	0.11513	0.14172	0.41480	0.13102	-0.04209	0.15877	-0.15521	0.02921	0.10360	1.00000	0.00164
Real Estate	0.00000	0.04924	0.22374	-0.00057	-0.10054	-0.08378	-0.09229	-0.17808	-0.20969	0.11344	-0.15282	-0.09055	0.00164	1.00000

Table 10 Correlation of original factors

	Mkt-RF	SMB	HML	Energy	Materials	Industrials	Cons Disc	Cons Stap	Healthcare	Financials	IT	Comm	Utilities	Real Estate
Mkt-RF	1.00000	0.06815	0.06374	-0.00597	-0.00967	-0.01271	-0.00548	0.02767	0.00210	0.00318	0.00145	0.02030	-0.01782	-0.01554
SMB	0.06815	1.00000	-0.18204	0.01706	0.02761	0.03630	0.01566	-0.07902	-0.00598	-0.00909	-0.00414	-0.05799	0.05090	0.04439
HML	0.06374	-0.18204	1.00000	0.01596	0.02582	0.03395	0.01464	-0.07390	-0.00560	-0.00850	-0.00387	-0.05423	0.04761	0.04151
Energy	-0.00597	0.01706	0.01596	1.00000	-0.00242	-0.00318	-0.00137	0.00693	0.00052	0.00080	0.00036	0.00508	-0.00446	-0.00389
Materials	-0.00967	0.02761	0.02582	-0.00242	1.00000	-0.00515	-0.00222	0.01121	0.00085	0.00129	0.00059	0.00822	-0.00722	-0.00630
Industrials	-0.01271	0.03630	0.03395	-0.00318	-0.00515	1.00000	-0.00292	0.01474	0.00112	0.00169	0.00077	0.01081	-0.00949	-0.00828
Cons Disc	-0.00548	0.01566	0.01464	-0.00137	-0.00222	-0.00292	1.00000	0.00636	0.00048	0.00073	0.00033	0.00466	-0.00409	-0.00357
Cons Stap	0.02767	-0.07902	-0.07390	0.00693	0.01121	0.01474	0.00636	1.00000	-0.00243	-0.00369	-0.00168	-0.02354	0.02067	0.01802
Healthcare	0.00210	-0.00598	-0.00560	0.00052	0.00085	0.00112	0.00048	-0.00243	1.00000	-0.00028	-0.00013	-0.00178	0.00156	0.00136
Financials	0.00318	-0.00909	-0.00850	0.00080	0.00129	0.00169	0.00073	-0.00369	-0.00028	1.00000	-0.00019	-0.00271	0.00238	0.00207
IT	0.00145	-0.00414	-0.00387	0.00036	0.00059	0.00077	0.00033	-0.00168	-0.00013	-0.00019	1.00000	-0.00123	0.00108	0.00094
Comm	0.02030	-0.05799	-0.05423	0.00508	0.00822	0.01081	0.00466	-0.02354	-0.00178	-0.00271	-0.00123	1.00000	0.01517	0.01322
Utilities	-0.01782	0.05090	0.04761	-0.00446	-0.00722	-0.00949	-0.00409	0.02067	0.00156	0.00238	0.00108	0.01517	1.00000	0.01161
Real Estate	-0.01554	0.04439	0.04151	-0.00389	-0.00630	-0.00828	-0.00357	0.01802	0.00136	0.00207	0.00094	0.01322	-0.01161	1.00000

Table 11 Correlation of orthogonalized factors using the GS method

Table 14 Correlation of original factors

Mkt-RF	SMB	HML	1010	1510	2010	2020	2030	2510	2520	2530	2550	3010	3020	3030	3510	3520	4010	4020	4030	4510	4520	4530	5010	5020	5510	6010
SMB	1.0000	0.06374	-0.00597	-0.00719	-0.01006	-0.01497	-0.00975	-0.00967	-0.00613	0.01747	-0.00152	0.02454	0.00018	-0.01113	0.00804	-0.04733	0.00388	0.01261	-0.00942	0.00134	0.02027	0.00107	0.02054	-0.01242	-0.01199	
HML	0.06374	1.0000	0.01586	0.02282	0.02028	0.02688	0.03998	0.02603	0.02186	-0.04668	0.00407	-0.06556	-0.00047	0.02973	0.02148	0.01265	-0.01037	-0.03368	-0.02515	-0.00559	-0.00553	-0.02027	-0.05486	0.03317	0.03204	
1010	-0.00597	0.01586	1.0000	0.00742	-0.00190	-0.00252	-0.00375	-0.00244	-0.00203	0.00438	-0.00308	0.00615	0.00004	-0.00279	0.00201	-0.00119	0.00097	0.00316	0.00034	0.00652	0.00027	0.00514	-0.00311	-0.00390	-0.00390	
1510	-0.00719	0.02282	0.00742	1.0000	-0.00408	-0.00498	-0.00686	-0.00595	-0.00329	0.00708	-0.00062	0.00954	0.00007	-0.00451	0.00157	0.00511	0.00381	0.00054	0.00384	0.00044	0.00352	-0.00503	-0.00586	0.02547	0.02426	
2010	-0.01006	0.02028	-0.00190	-0.00408	1.0000	-0.00476	-0.00310	-0.00258	0.00004	0.00556	-0.00049	0.00781	0.00006	-0.00354	0.00256	-0.00151	0.00123	0.00401	0.00300	0.00043	0.00066	0.00034	0.00653	-0.00395	-0.00382	
2020	-0.00975	0.02688	-0.00252	-0.00498	-0.00476	1.0000	-0.00351	-0.00261	0.00000	0.00737	-0.00064	0.01135	0.00007	-0.00469	0.00339	-0.00200	0.00184	0.00532	0.00087	0.00087	0.00045	0.00886	-0.00524	-0.00506		
2030	-0.01497	0.03998	-0.00375	-0.00686	-0.00310	-0.00261	1.0000	-0.00509	-0.00098	0.01096	-0.00096	0.00911	-0.00688	0.00564	-0.00297	0.00243	0.00791	0.00591	0.00084	0.00130	0.00067	0.01388	-0.00779	-0.00752		
2510	-0.00967	0.02603	-0.00244	-0.00595	-0.00310	-0.00411	-0.00509	1.0000	-0.00331	0.00574	-0.00062	0.01022	0.00007	-0.00455	0.00328	-0.00193	0.00158	0.00515	0.00385	0.00055	0.00044	0.00839	-0.00507	-0.00490		
2520	-0.00613	0.02186	-0.00203	-0.00375	-0.00244	-0.00261	-0.00509	-0.00331	1.0000	0.00094	0.00006	0.00378	0.00273	-0.00161	0.00132	0.00429	0.00220	0.00046	0.00070	0.00070	0.00037	0.00698	-0.00422	-0.00408		
2530	-0.00063	0.02186	-0.00203	-0.00375	-0.00244	-0.00261	-0.00509	-0.00331	0.00094	1.0000	0.00010	0.00112	-0.01798	-0.00013	0.00815	-0.00589	0.00347	-0.00284	-0.00923	-0.00089	-0.00098	-0.00152	-0.00079	-0.01504	0.00909	
2550	0.01747	-0.04668	0.00407	0.00708	0.00556	0.00737	0.01096	0.00714	0.00594	0.00010	1.0000	0.00112	-0.01798	-0.00013	0.00815	-0.00589	0.00347	-0.00284	-0.00923	-0.00089	-0.00098	-0.00152	-0.00079	-0.01504	0.00909	
3010	-0.00152	0.00407	-0.00064	-0.00096	-0.00049	-0.00064	-0.00096	-0.00062	-0.00049	0.00112	0.00112	1.0000	0.00112	-0.01798	-0.00013	0.00815	-0.00589	0.00347	-0.00284	-0.00923	-0.00089	-0.00098	-0.00152	-0.00079		
3020	0.02454	-0.00047	0.00004	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00112	0.00112	0.00112	1.0000	0.00112	-0.01798	-0.00013	0.00815	-0.00589	0.00347	-0.00284	-0.00923	-0.00089	-0.00098	-0.00152		
3030	0.00018	-0.00047	0.00004	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00112	0.00112	0.00112	0.00112	1.0000	0.00112	-0.01798	-0.00013	0.00815	-0.00589	0.00347	-0.00284	-0.00923	-0.00089	-0.00098		
3510	-0.01113	0.00804	-0.00279	0.00201	-0.00119	0.00097	0.00316	0.00034	0.00652	0.00007	0.00514	-0.00311	0.00000	0.00000	1.0000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000		
3520	0.00804	-0.04733	0.00388	0.01261	-0.00942	0.00134	0.02027	0.00107	0.02054	-0.01242	-0.01199	0.00000	0.00000	0.00000	0.00000	1.0000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000		
4010	-0.04733	0.01261	-0.00942	0.00134	0.02027	0.00107	0.02054	-0.01242	-0.01199	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	1.0000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000		
4020	-0.00388	0.01261	-0.00942	0.00134	0.02027	0.00107	0.02054	-0.01242	-0.01199	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	1.0000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000		
4030	-0.00134	0.02027	0.00107	0.02054	-0.01242	-0.01199	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	1.0000	0.00000	0.00000	0.00000	0.00000	0.00000		
4510	0.00942	-0.00942	0.00055	0.00055	0.00055	0.00055	0.00055	0.00055	0.00055	0.00055	0.00055	0.00055	0.00055	0.00055	0.00055	0.00055	0.00055	0.00055	0.00055	1.0000	0.00000	0.00000	0.00000	0.00000		
4520	0.00134	-0.00942	0.00055	0.00055	0.00055	0.00055	0.00055	0.00055	0.00055	0.00055	0.00055	0.00055	0.00055	0.00055	0.00055	0.00055	0.00055	0.00055	0.00055	0.00055	1.0000	0.00000	0.00000	0.00000		
4530	0.00007	-0.00942	0.00055	0.00055	0.00055	0.00055	0.00055	0.00055	0.00055	0.00055	0.00055	0.00055	0.00055	0.00055	0.00055	0.00055	0.00055	0.00055	0.00055	0.00055	0.00055	1.0000	0.00000	0.00000		
5010	0.00107	-0.00300	0.00043	0.00066	0.00034	0.00653	-0.00395	-0.00382	-0.00390	-0.00390	-0.00390	-0.00390	-0.00390	-0.00390	-0.00390	-0.00390	-0.00390	-0.00390	-0.00390	-0.00390	-0.00390	-0.00390	1.0000	0.00000		
5020	0.00034	0.00066	0.00034	0.00653	-0.00395	-0.00382	-0.00390	-0.00390	-0.00390	-0.00390	-0.00390	-0.00390	-0.00390	-0.00390	-0.00390	-0.00390	-0.00390	-0.00390	-0.00390	-0.00390	-0.00390	-0.00390	-0.00390	1.0000		
5030	0.00066	0.00034	0.00653	-0.00395	-0.00382	-0.00390	-0.00390	-0.00390	-0.00390	-0.00390	-0.00390	-0.00390	-0.00390	-0.00390	-0.00390	-0.00390	-0.00390	-0.00390	-0.00390	-0.00390	-0.00390	-0.00390	-0.00390	-0.00390		
5510	-0.01242	-0.01199	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000		
6010	-0.01199	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000		

Table 15 Correlation of orthogonalized factors using the GS method

Mkt-RF	SMB	HML	1010	1510	2010	2020	2030	2510	2520	2530	2550	3010	3020	3030	3510	3520	4010	4020	4030	4510	4520	4530	5010	5020	5510	6010
SMB	1.0000	0.06374	-0.00597	-0.00719	-0.01006	-0.01497	-0.00975	-0.00967	-0.00613	0.01747	-0.00152	0.02454	0.00018	-0.01113	0.00804	-0.04733	0.00388	0.01261	-0.00942	0.00134	0.02027	0.00107	0.02054	-0.01242	-0.01199	
HML	0.06374	1.0000	0.01586	0.02282	0.02028	0.02688	0.03998	0.02603	0.02186	-0.04668	0.00407	-0.06556	-0.00047	0.02973	0.02148	0.01265	-0.01037	-0.03368	-0.02515	-0.00559	-0.00553	-0.02027	-0.05486	0.03317	0.03204	
1010	-0.00597	0.01586	1.0000	0.00742	-0.00190	-0.00252	-0.00375	-0.00244	-0.00203	0.00438	-0.00308	0.00615	0.00004	-0.00279	0.00201	-0.00119	0.00097	0.00316	0.00034	0.00652	0.00027	0.00514	-0.00311	-0.00390		
1510	-0.00719	0.02282	0.00742	1.0000	-0.00408	-0.00498	-0.00686	-0.00595	-0.00329	0.00708	-0.00062	0.00954	0.00007	-0.00451	0.00328	-0.00193	0.00158	0.00515	0.00385	0.00055	0.00044	0.00839	-0.00507	-0.00490		
2010	-0.01006	0.02028	-0.00190	-0.00408	1.0000	-0.00476	-0.00310	-0.00258	0.00004	0.00556	-0.00049	0.00781	0.00006	-0.00354	0.00256	-0.00151	0.00123	0.00401	0.00300	0.00043	0.00066	0.00034	0.00653	-0.00395		
2020	-0.00975	0.02688	-0.00252	-0.00498	-0.00476	1.0000	-0.00351	-0.00261	0.00000	0.00737	-0.00064	0.01135	0.00007	-0.00469	0.00339	-0.00200	0.00184	0.00532	0.00087	0.00087	0.00045	0.00886	-0.00524	-0.00506		
2030	-0.01497	0.03998	-0.00375	-0.00686	-0.00310	-0.00261	1.0000	-0.00509	-0.00098	0.01096	-0.00096	0.00911	-0.00688	0.00564	-0.00297	0.00243	0.00791	0.00591	0.00084	0.00130	0.00067	0.01388	-0.00779	-0.00752		
2510	-0.00967	0.02603	-0.00244	-0.00595	-0.00310	-0.00411	-0.00509	1.0000	-0.00331	0.00574	-0.00062	0.01022	0.00007	-0.00455	0.00328	-0.00193	0.00158	0.00515	0.00385	0.00055	0.00044	0.00839	-0.00507	-0.00490		
2520	-0.00613	0.02186	-0.00203	-0.00																						

stability is not as high as in the US market. This might be due to the smaller universe we use to construct the risk model for the Chinese market. On average 67.29% of the variance can be explained by the model. This is higher than what can be achieved for the corresponding risk model in the US market.

5. Conclusion

The results we obtained from the above four models demonstrate that the SW/L method gives a significantly higher overlap between the orthogonalized factors and the original factors than what can be obtained using the GS method, especially when more factors are included in the risk model. The coefficient matrix relating old factors to new factors is also shown to be quite stable; this allows for the construction of a robust risk model based on orthogonalized factors which can retain the economic intuition contained in the original factors. Our 27-factor risk model can be used as a baseline model, as it can explain a significant percentage of variance in the stock returns. Our pilot study of this model points to the feasibility of the SW/L method for constructing a useful risk model with linearly independent factors. We plan to extend our study to construct a more complete risk model by using additional style factors and a finer level of the industry exposure (namely exposures to 69 GICS industries) and using a bigger universe.

Reference

- [1] Fama, E. F.; French, K. R. (1993). "Common risk factors in the returns on stocks and bonds". *Journal of Financial Economics*. 33: 3–56.
- [2] Jegadeesh, N., & Titman, S. . (1993). Returns to buying winners and selling losers: implications for stock market efficiency. *Journal of Finance*, 48(1), 65-91.
- [3] Chen, H. L., & Bondt, W. D. . (2004). Style momentum within the s&p-500 index. *Journal of Empirical Finance*, 11(4), 483-507.
- [4] Löwdin, P.-O. (1950). On the non-orthogonality problem connected with the use of atomic wave functions in the theory of molecules and crystals. *The Journal of Chemical Physics*, 18(3), 365–375
- [5] Schweinler, H.C., Wigner, E.P. (1970). Orthogonalization methods. *Mathematical Physics*, 11, 1693–1694.
- [6] Klein, R.F., Chow, V.K. (2013). Orthogonalized factors and systematic risk decomposition. *Quarterly Review of Economics and Finance*, 53(2), 175–187.
- [7] Klein, R. F. . (2011). Analysis of systematic risk: decomposition and portfolio efficiency. *Dissertations & Theses - Gradworks*.
- [8] Adcock, C., Bessler, W., & Conlon, T. . (2019). Fundamental factor models and macroeconomic risks - an orthogonal decomposition. *SSRN Electronic Journal*.