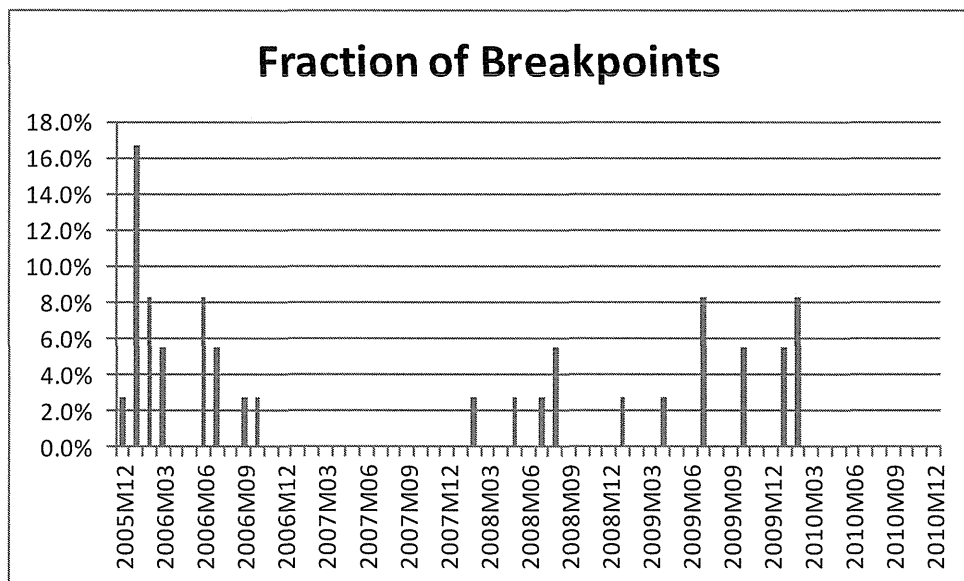


Figure 5 Break Point of Style Based on Wald F-statistic with p-value less than 0.1 (i.e. significance level 10%)



Appendix A

Principal Component Analysis

The model that is applied to the data is principal component analysis. This approach is used when the fund returns are highly correlated and may be represented by several primary underlying unobservable stochastic factors. For an introduction to principal component analysis, one may refer to any basic statistics textbook such as Brooks (2002). Because many of the fund are focusing on socially responsible investments their returns are highly correlated. In such a situation, the returns are very likely to be explained by a handful of dominant stochastic factors when expressed in a system of linear equations. This is one motivation to analyze the funds using principal component analysis. Following the notation of Brooks (2002), we have a system of linear equations,

$$\begin{aligned} p_1 &= \alpha_{11}x_1 + \alpha_{12}x_2 + \cdots + \alpha_{1k}x_k \\ p_2 &= \alpha_{21}x_1 + \alpha_{22}x_2 + \cdots + \alpha_{2k}x_k \\ &\dots\dots\dots \\ p_k &= \alpha_{k1}x_1 + \alpha_{k2}x_2 + \cdots + \alpha_{kk}x_k \end{aligned}$$

where α_{ij} are the coefficients of explanatory variable j for principal component i . There are a total of k principal components $p_1 \cdots p_k$ and $x_1 \cdots x_k$ that correspond to the explanatory variables. In matrix notation, the system of equations may be represented as

$$p = Ax$$

where

$$p = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_k \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix}, \quad A = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1k} \\ \alpha_{21} & \ddots & & \alpha_{2k} \\ \vdots & & \ddots & \vdots \\ \alpha_{k1} & \cdots & \cdots & \alpha_{kk} \end{bmatrix}.$$

Or equivalently, in sigma notation, it may be expressed as

$$p_i = \sum_{j=1}^k \alpha_{ij}x_j \quad \forall i = 1, \dots, k$$

The coefficients α_{ij} are restricted in the sense that the square of the row elements must sum up to unity. If the number of observations used to estimate the coefficients is T , then each explanatory variable will have T observations that can be represented as the following.

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1T} \\ x_{21} & \ddots & & x_{2T} \\ \vdots & & \ddots & \vdots \\ x_{k1} & \cdots & \cdots & x_{kT} \end{bmatrix}$$

Here, each row corresponds to the array of observations for an explanatory variable. In particular, these variables correspond to the returns of the SC funds. There are k funds in total, where the fund returns are calculated over a period with T observations.⁷ The principal components may be interpreted as the eigenvalues of the design matrix XX' . The number of eigenvalues will be equal to the rank of the design matrix $rank(XX')$, which is k because the set of linear equations is not degenerate. Denoting the eigenvalues as λ_i for $i=1, \dots, k$, the proportion of variance that is explained by this component can be expressed as the following.

$$\phi_i = \frac{\lambda_i}{\sum_{i=1}^k \lambda_i} \quad \text{for all } i=1, \dots, k$$

Where ϕ_i is the proportion that is explained by principal component i .

Appendix B

Description of Apparel & luxury goods Index

The index is a sub-index of the SP500 index with an industry code 252030 under the name “Textiles, Apparel & Luxury Goods”. This index is comprised of the following 3 sub-industries.

25203010	Apparel, Accessories & Luxury Goods
	Manufacturers of apparel, accessories & luxury goods. Includes companies primarily producing designer handbags, wallets, luggage, jewelry and watches. Excludes shoes classified in the Footwear Sub-Industry.
25203020	Footwear
	Manufacturers of footwear. Includes sport and leather shoes.
25203030	Textiles
	Manufacturers of textile and related products not classified in the Apparel, Accessories & Luxury Goods, Footwear or Home Furnishings Sub-Industries.

The Industry group is 2520 Consumer Durables & Apparel, and industry sector is 25 Consumer Discretionary.

⁷ The analysis was conducted with monthly returns (not annualized) for 7 funds. There were 72 observations in total (starting from 2005 Jan. to 2010 Dec.)

Appendix C

Rolling Window Style Regressions

The rolling style regression analysis in this paper is very similar to the model of Sharpe (1992). The model minimizes the squared residuals of the regression, and restricts the coefficients to be between 0 and 1. The estimation process was conducted by solving the following optimization problem.

$$\begin{aligned} & \text{Min } \varepsilon^T \varepsilon \\ \text{s. t. } & \varepsilon = Y - (\alpha_1 X_1 - \alpha_2 X_2) \\ & \alpha_1 = (t_1)^2 \\ & \alpha_2 = (t_2)^2 \\ & (t_1)^2 + (t_2)^2 = 1 \\ & t_1, t_2 \geq 0 \end{aligned}$$

where ε is the residual of the regression, Y is the overall equity market approximated by the TOPIX index, X_1 is the returns of the apparel & luxury goods index (sub index of SP500 index), X_2 is the returns of the socially conscious fund index, and α_1, α_2 are the weights associated to each of the dependent variables. The coefficients are restricted to be between 0 and 1, and estimated by minimizing the sum squared of errors.

The model yields coefficient estimates comprised of two factors, the Japanese market index factor and the luxury consumption factor. The explanatory variable is the SC fund which is expressed by the two factors with a non-negative weight between 0 and 1.

Appendix D Descriptive Statistics

Socially Conscious Funds

Fund Number/Name	notation
F2/AsahiLife SRI	F2_RET
F4/Daiwa DC SRI	F4_RET
F5/Daiwa SRI	F5_RET
F6/MitsubishiUFJ SRI	F6_RET
F9/Shinkin Fukoku SRI	F9_RET
F12/STAM DC Good Company SRI	F12_RET
F14/STAM SRI Japan Open	F14_RET

	F2_RET	F4_RET	F5_RET	F6_RET	F12_RET	F14_RET
Mean	-0.00	-0.00	-0.00	-0.00	-0.00	-0.01
Median	0.00	0.00	0.00	0.00	-0.00	-0.00
Maximum	0.12	0.13	0.13	0.13	0.15	0.16
Minimum	-0.20	-0.21	-0.23	-0.22	-0.22	-0.27
Std. Dev.	0.05	0.06	0.06	0.06	0.06	0.07
Skewness	-0.80	-0.60	-1.03	-0.57	-0.51	-1.04
Kurtosis	4.64	4.39	5.44	4.61	4.51	6.12

TOPIX and its 33 sectors

notation	SECTOR
T1	Topix Air Transport PR JPY
T2	Topix Banks PR JPY
T3	Topix Chemicals PR JPY
T4	Topix Communication PR JPY
T5	Topix Construction PR JPY
T6	Topix Electronic Appliances PR JPY
T7	Topix Electronic Power&Gas PR JPY
T8	Topix Financing Business PR JPY
T9	Topix Fishery Agricul&Forest PR JPY
T10	Topix Foods PR JPY
T11	Topix Glass&Ceramics PR JPY
T12	Topix Insurance PR JPY
T13	Topix Iron&Steel PR JPY
T14	Topix Land Transport PR JPY
T15	Topix Machinery PR JPY
T16	Topix Marine Transport PR JPY
T17	Topix Metal Products PR JPY
T18	Topix Mining PR JPY
T19	Topix NonFerrous Metals PR JPY
T20	Topix Oil&Coal Products PR JPY
T21	Topix Other Products PR JPY
T22	Topix Pharmaceutical PR JPY
T23	Topix PR JPY
T24	Topix Precision Instruments PR JPY
T25	Topix Pulp&Paper PR JPY
T26	Topix Real Estate PR JPY
T27	Topix Retail Trade PR JPY
T28	Topix Rubber Products PR JPY
T29	Topix Securities PR JPY
T30	Topix Services PR JPY
T31	Topix Textiles&Apparels PR JPY
T32	Topix Transport Equipment PR JPY
T33	Topix Warehousing PR JPY
T34	Topix Wholesale Trade PR JPY

	Mean	Median	Maximum	Minimum	Std. Dev.	Skewness	Kurtosis
T1	-0.01	-0.01	0.13	-0.18	0.06	-0.42	4.23
T2	-0.01	-0.01	0.28	-0.30	0.08	0.15	6.65
T3	0.00	0.00	0.11	-0.16	0.06	-0.50	3.33
T4	-0.00	0.00	0.09	-0.15	0.05	-0.52	3.34
T5	-0.00	-0.01	0.14	-0.13	0.06	0.09	2.24
T6	0.00	0.00	0.17	-0.25	0.07	-0.41	4.36
T7	-0.00	-0.00	0.08	-0.12	0.04	-0.67	3.55
T8	-0.01	-0.01	0.28	-0.31	0.10	0.00	3.88
T9	-0.00	0.01	0.15	-0.16	0.05	-0.36	3.67
T10	0.00	0.01	0.10	-0.15	0.04	-0.86	4.73
T11	0.00	0.01	0.15	-0.29	0.08	-0.60	4.26
T12	-0.00	-0.00	0.23	-0.23	0.08	0.15	3.95
T13	0.01	0.00	0.27	-0.27	0.10	0.11	3.71
T14	-0.00	-0.00	0.10	-0.10	0.04	-0.04	3.21
T15	0.01	0.01	0.15	-0.29	0.08	-0.86	4.56
T16	0.00	0.01	0.18	-0.38	0.10	-0.94	4.79
T17	-0.00	-0.00	0.12	-0.16	0.06	-0.19	2.64
T18	0.00	0.01	0.24	-0.37	0.10	-0.67	4.73
T19	0.00	0.01	0.19	-0.32	0.09	-0.70	4.80
T20	0.00	0.00	0.15	-0.24	0.09	-0.46	3.10
T21	0.00	0.01	0.20	-0.24	0.07	-0.56	4.69
T22	-0.00	0.00	0.08	-0.15	0.05	-0.64	3.53
T23	-0.00	0.00	0.12	-0.20	0.06	-0.61	4.37
T24	0.00	0.00	0.13	-0.29	0.07	-0.98	5.41
T25	-0.00	-0.01	0.21	-0.23	0.07	-0.13	5.15
T26	0.00	0.02	0.27	-0.19	0.10	0.26	3.20
T27	-0.00	-0.00	0.14	-0.14	0.06	-0.18	2.96
T28	-0.00	0.00	0.18	-0.18	0.07	-0.09	2.84
T29	-0.01	-0.01	0.21	-0.28	0.11	-0.15	3.14
T30	-0.00	0.00	0.09	-0.16	0.05	-0.75	3.80
T31	-0.00	-0.00	0.15	-0.17	0.06	-0.04	3.25
T32	0.00	0.01	0.23	-0.21	0.07	0.04	4.12
T33	-0.00	-0.01	0.19	-0.14	0.06	0.32	3.14
T34	0.01	0.00	0.16	-0.27	0.08	-0.67	4.28

Table 7 Correlation matrix of TOPIX indices

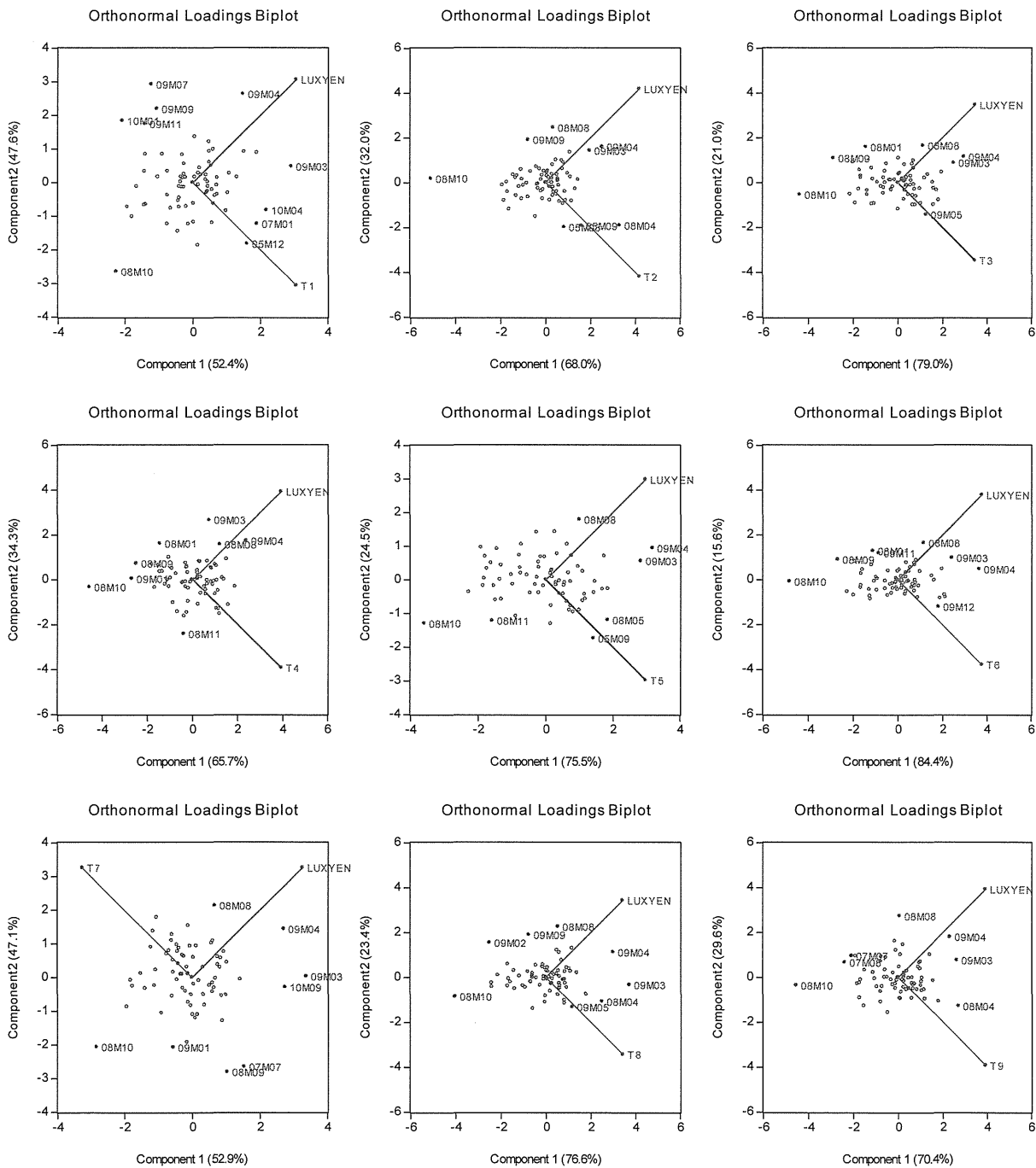
	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14	T15	T16	T17
T1	1.00	0.11	0.21	0.16	0.31	0.12	0.08	0.23	0.09	0.12	0.20	0.12	0.08	0.32	0.16	0.17	0.27
T2	0.11	1.00	0.67	0.59	0.72	0.62	0.10	0.78	0.57	0.51	0.63	0.80	0.63	0.55	0.70	0.61	0.64
T3	0.21	0.67	1.00	0.67	0.82	0.91	0.14	0.70	0.58	0.72	0.91	0.66	0.81	0.61	0.92	0.72	0.87
T4	0.16	0.59	0.67	1.00	0.66	0.56	0.43	0.50	0.47	0.68	0.56	0.50	0.56	0.70	0.65	0.62	0.68
T5	0.31	0.72	0.82	0.66	1.00	0.70	0.14	0.76	0.55	0.58	0.77	0.69	0.70	0.63	0.79	0.63	0.82
T6	0.12	0.62	0.91	0.56	0.70	1.00	0.03	0.64	0.57	0.62	0.89	0.64	0.75	0.49	0.87	0.69	0.82
T7	0.08	0.10	0.14	0.43	0.14	0.03	1.00	0.03	0.21	0.36	0.00	0.08	0.17	0.45	0.04	0.12	0.22
T8	0.23	0.78	0.70	0.50	0.76	0.64	0.03	1.00	0.61	0.53	0.66	0.67	0.61	0.45	0.66	0.55	0.67
T9	0.09	0.57	0.58	0.47	0.55	0.57	0.21	0.61	1.00	0.55	0.55	0.45	0.58	0.38	0.59	0.49	0.61
T10	0.12	0.51	0.72	0.68	0.58	0.62	0.36	0.53	0.55	1.00	0.64	0.45	0.62	0.58	0.72	0.53	0.69
T11	0.20	0.63	0.91	0.56	0.77	0.89	0.00	0.66	0.55	0.64	1.00	0.59	0.75	0.54	0.89	0.69	0.82
T12	0.12	0.80	0.66	0.50	0.69	0.64	0.08	0.67	0.45	0.45	0.59	1.00	0.63	0.55	0.66	0.57	0.60
T13	0.08	0.63	0.81	0.56	0.70	0.75	0.17	0.61	0.58	0.62	0.75	0.63	1.00	0.47	0.81	0.71	0.68
T14	0.32	0.55	0.61	0.70	0.63	0.49	0.45	0.45	0.38	0.58	0.54	0.55	0.47	1.00	0.57	0.50	0.61
T15	0.16	0.70	0.92	0.65	0.79	0.87	0.04	0.66	0.59	0.72	0.89	0.66	0.81	0.57	1.00	0.78	0.84
T16	0.17	0.61	0.72	0.62	0.63	0.69	0.12	0.55	0.49	0.53	0.69	0.57	0.71	0.50	0.78	1.00	0.63
T17	0.27	0.64	0.87	0.68	0.82	0.82	0.22	0.67	0.61	0.69	0.82	0.60	0.68	0.61	0.84	0.63	1.00
T18	0.15	0.46	0.66	0.54	0.49	0.57	0.08	0.34	0.35	0.50	0.62	0.41	0.53	0.41	0.70	0.70	0.51
T19	0.15	0.67	0.87	0.61	0.72	0.82	0.02	0.60	0.52	0.60	0.84	0.62	0.78	0.60	0.91	0.79	0.72
T20	0.22	0.55	0.71	0.58	0.62	0.66	-0.00	0.44	0.46	0.47	0.68	0.60	0.67	0.40	0.74	0.70	0.62
T21	0.10	0.58	0.74	0.52	0.58	0.68	0.09	0.55	0.43	0.57	0.75	0.48	0.57	0.45	0.75	0.73	0.63
T22	0.03	0.56	0.67	0.71	0.61	0.62	0.43	0.46	0.58	0.70	0.55	0.49	0.66	0.54	0.62	0.62	0.63
T23	0.19	0.82	0.95	0.73	0.84	0.91	0.18	0.76	0.64	0.72	0.88	0.77	0.84	0.67	0.93	0.78	0.86
T24	0.06	0.58	0.87	0.56	0.63	0.86	0.03	0.57	0.50	0.70	0.86	0.54	0.69	0.43	0.82	0.60	0.75
T25	0.17	0.49	0.52	0.64	0.54	0.45	0.27	0.46	0.40	0.55	0.44	0.45	0.47	0.53	0.58	0.48	0.57
T26	0.35	0.75	0.75	0.62	0.79	0.65	0.11	0.81	0.56	0.58	0.67	0.72	0.66	0.63	0.70	0.58	0.70
T27	0.25	0.74	0.79	0.77	0.79	0.68	0.23	0.71	0.58	0.67	0.69	0.64	0.60	0.73	0.74	0.55	0.76
T28	0.15	0.39	0.62	0.30	0.46	0.60	-0.01	0.39	0.45	0.47	0.60	0.38	0.52	0.35	0.62	0.39	0.53
T29	0.20	0.80	0.77	0.58	0.77	0.72	0.10	0.88	0.56	0.58	0.76	0.67	0.69	0.57	0.73	0.60	0.70
T30	0.28	0.62	0.82	0.80	0.76	0.71	0.25	0.60	0.56	0.77	0.76	0.55	0.59	0.76	0.82	0.61	0.81
T31	0.27	0.64	0.83	0.62	0.78	0.78	0.07	0.68	0.56	0.65	0.79	0.56	0.68	0.65	0.86	0.61	0.79
T32	0.05	0.53	0.76	0.42	0.57	0.84	0.04	0.50	0.41	0.48	0.77	0.57	0.74	0.39	0.73	0.57	0.61
T33	0.41	0.59	0.70	0.67	0.79	0.52	0.23	0.63	0.57	0.64	0.63	0.57	0.52	0.71	0.68	0.53	0.74
T34	0.15	0.62	0.86	0.56	0.68	0.79	0.01	0.62	0.50	0.56	0.82	0.62	0.83	0.50	0.85	0.78	0.68

Table 7 Correlation matrix of TOPIX indices

	T18	T19	T20	T21	T22	T23	T24	T25	T26	T27	T28	T29	T30	T31	T32	T33	T34
T1	0.15	0.15	0.22	0.10	0.03	0.19	0.06	0.17	0.35	0.25	0.15	0.20	0.28	0.27	0.05	0.41	0.15
T2	0.46	0.67	0.55	0.58	0.56	0.82	0.58	0.49	0.75	0.74	0.39	0.80	0.62	0.64	0.53	0.59	0.62
T3	0.66	0.87	0.71	0.74	0.67	0.95	0.87	0.52	0.75	0.79	0.62	0.77	0.82	0.83	0.76	0.70	0.86
T4	0.54	0.61	0.58	0.52	0.71	0.73	0.56	0.64	0.62	0.77	0.30	0.58	0.80	0.62	0.42	0.67	0.56
T5	0.49	0.72	0.62	0.58	0.61	0.84	0.63	0.54	0.79	0.79	0.46	0.77	0.76	0.78	0.57	0.79	0.68
T6	0.57	0.82	0.66	0.68	0.62	0.91	0.86	0.45	0.65	0.68	0.60	0.72	0.71	0.78	0.84	0.52	0.79
T7	0.08	0.02	-0.00	0.09	0.43	0.18	0.03	0.27	0.11	0.23	-0.01	0.10	0.25	0.07	0.04	0.23	0.01
T8	0.34	0.60	0.44	0.55	0.46	0.76	0.57	0.46	0.81	0.71	0.39	0.88	0.60	0.68	0.50	0.63	0.62
T9	0.35	0.52	0.46	0.43	0.58	0.64	0.50	0.40	0.56	0.58	0.45	0.56	0.56	0.56	0.41	0.57	0.50
T10	0.50	0.60	0.47	0.57	0.70	0.72	0.70	0.55	0.58	0.67	0.47	0.58	0.77	0.65	0.48	0.64	0.56
T11	0.62	0.84	0.68	0.75	0.55	0.88	0.86	0.44	0.67	0.69	0.60	0.76	0.76	0.79	0.77	0.63	0.82
T12	0.41	0.62	0.60	0.48	0.49	0.77	0.54	0.45	0.72	0.64	0.38	0.67	0.55	0.56	0.57	0.57	0.62
T13	0.53	0.78	0.67	0.57	0.66	0.84	0.69	0.47	0.66	0.60	0.52	0.69	0.59	0.68	0.74	0.52	0.83
T14	0.41	0.60	0.40	0.45	0.54	0.67	0.43	0.53	0.63	0.73	0.35	0.57	0.76	0.65	0.39	0.71	0.50
T15	0.70	0.91	0.74	0.75	0.62	0.93	0.82	0.58	0.70	0.74	0.62	0.73	0.82	0.86	0.73	0.68	0.85
T16	0.70	0.79	0.70	0.73	0.62	0.78	0.60	0.48	0.58	0.55	0.39	0.60	0.61	0.61	0.57	0.53	0.78
T17	0.51	0.72	0.62	0.63	0.63	0.86	0.75	0.57	0.70	0.76	0.53	0.70	0.81	0.79	0.61	0.74	0.68
T18	1.00	0.73	0.71	0.64	0.51	0.65	0.58	0.49	0.40	0.53	0.23	0.49	0.60	0.52	0.44	0.43	0.73
T19	0.73	1.00	0.75	0.68	0.61	0.89	0.76	0.48	0.68	0.68	0.56	0.71	0.74	0.76	0.73	0.61	0.89
T20	0.71	0.75	1.00	0.56	0.56	0.73	0.62	0.47	0.54	0.56	0.46	0.49	0.62	0.56	0.57	0.48	0.74
T21	0.64	0.68	0.56	1.00	0.44	0.74	0.71	0.46	0.51	0.62	0.40	0.64	0.62	0.62	0.56	0.54	0.68
T22	0.51	0.61	0.56	0.44	1.00	0.72	0.59	0.51	0.55	0.57	0.42	0.53	0.62	0.54	0.51	0.48	0.57
T23	0.65	0.89	0.73	0.74	0.72	1.00	0.84	0.57	0.82	0.83	0.60	0.84	0.84	0.84	0.80	0.71	0.87
T24	0.58	0.76	0.62	0.71	0.59	0.84	1.00	0.43	0.61	0.65	0.63	0.66	0.71	0.68	0.75	0.57	0.74
T25	0.49	0.48	0.47	0.46	0.51	0.57	0.43	1.00	0.41	0.61	0.18	0.50	0.65	0.59	0.26	0.51	0.37
T26	0.40	0.68	0.54	0.51	0.55	0.82	0.61	0.41	1.00	0.76	0.47	0.79	0.71	0.69	0.56	0.76	0.68
T27	0.53	0.68	0.56	0.62	0.57	0.83	0.65	0.61	0.76	1.00	0.50	0.78	0.85	0.81	0.51	0.80	0.63
T28	0.23	0.56	0.46	0.40	0.42	0.60	0.63	0.18	0.47	0.50	1.00	0.43	0.56	0.61	0.64	0.47	0.46
T29	0.49	0.71	0.49	0.64	0.53	0.84	0.66	0.50	0.79	0.78	0.43	1.00	0.69	0.71	0.61	0.64	0.71
T30	0.60	0.74	0.62	0.62	0.62	0.84	0.71	0.65	0.71	0.85	0.56	0.69	1.00	0.82	0.57	0.80	0.65
T31	0.52	0.76	0.56	0.62	0.54	0.84	0.68	0.59	0.69	0.81	0.61	0.71	0.82	1.00	0.65	0.74	0.69
T32	0.44	0.73	0.57	0.56	0.51	0.80	0.75	0.26	0.56	0.51	0.64	0.61	0.57	0.65	1.00	0.37	0.74
T33	0.43	0.61	0.48	0.54	0.48	0.71	0.57	0.51	0.76	0.80	0.47	0.64	0.80	0.74	0.37	1.00	0.55
T34	0.73	0.89	0.74	0.68	0.57	0.87	0.74	0.37	0.68	0.63	0.46	0.71	0.65	0.69	0.74	0.55	1.00

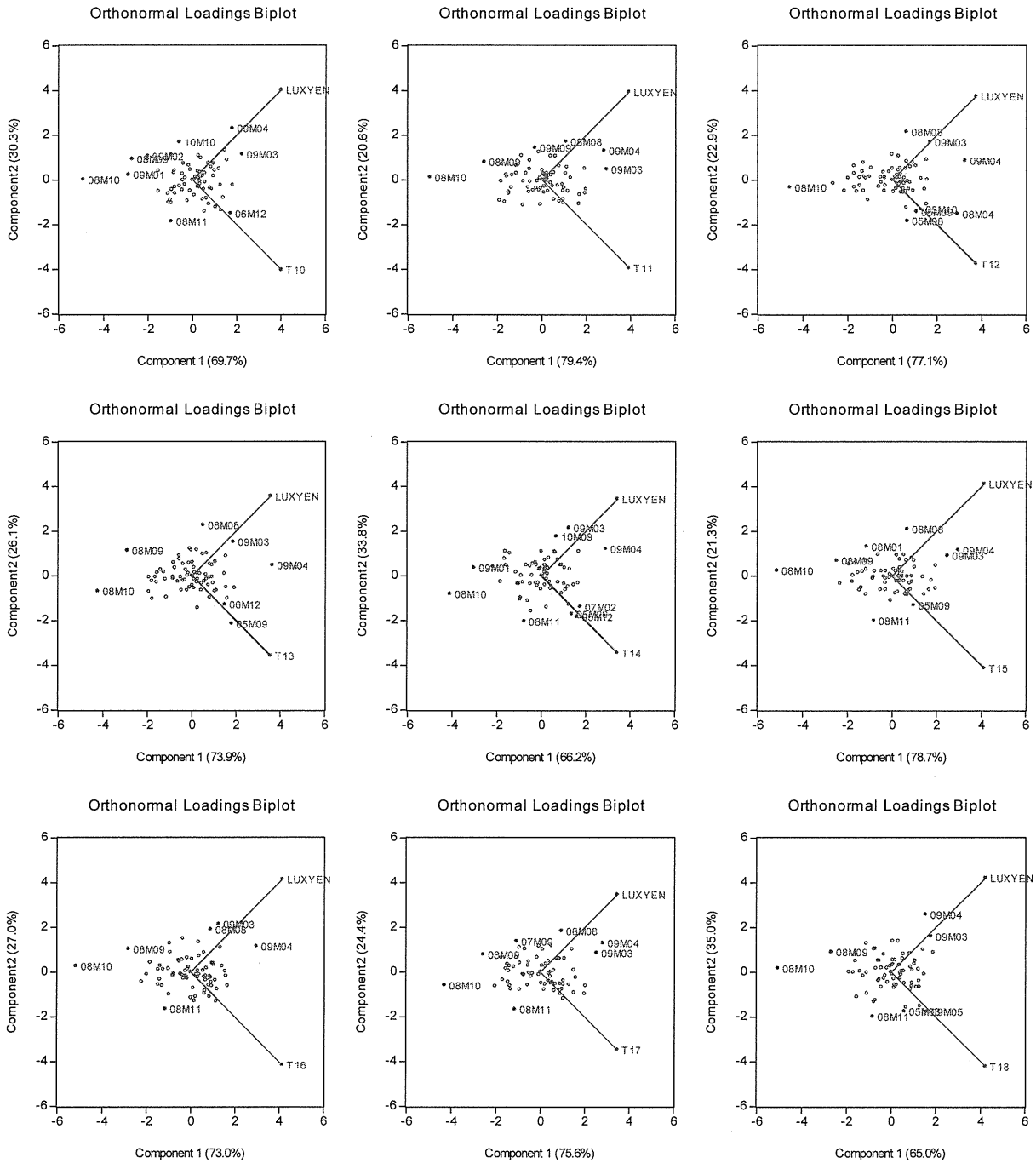
Appendix E

Orthonormal biplots of Apparel & Luxury Goods Index (yen base) versus TOPIX and its 33 sectors



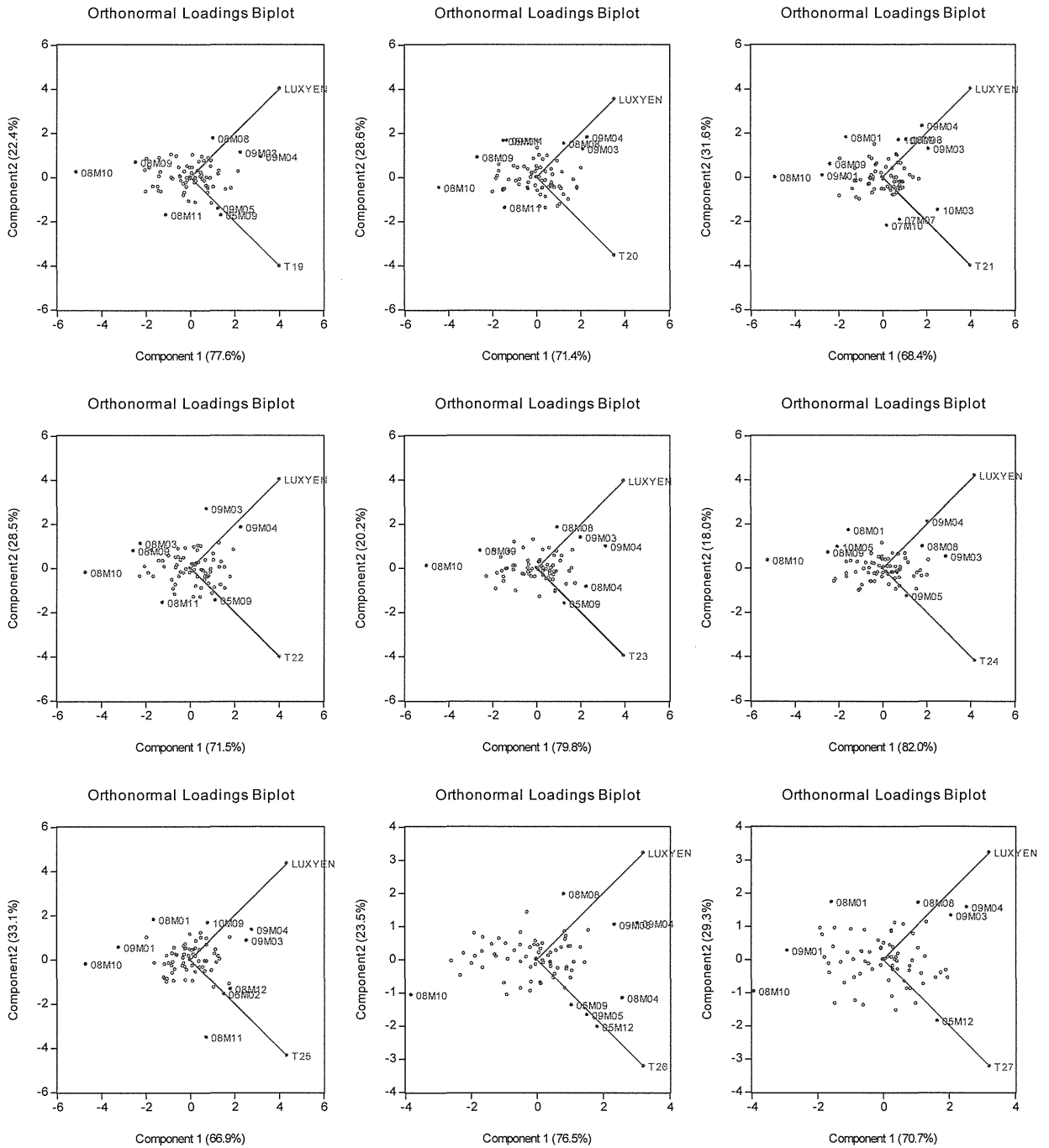
Appendix E (continued)

Orthonormal biplots of Apparel & Luxury Goods Index (yen base) versus TOPIX and its 33 sectors



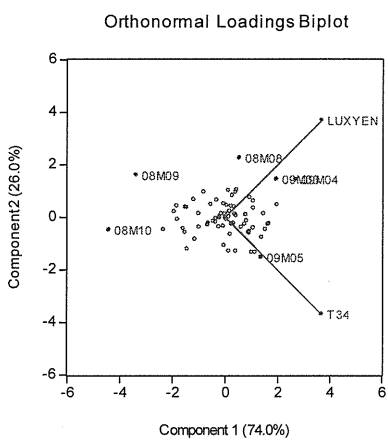
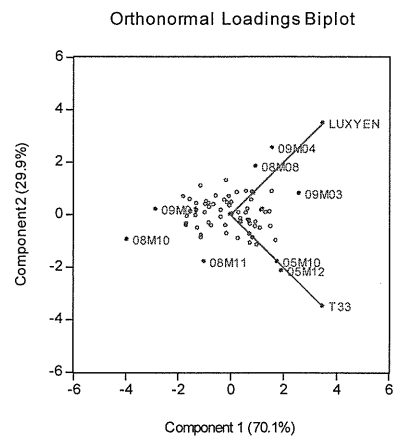
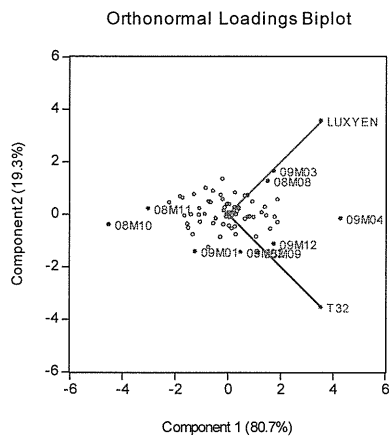
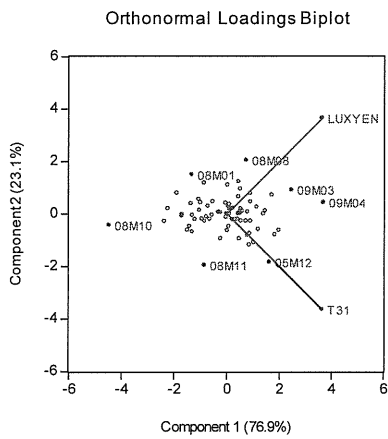
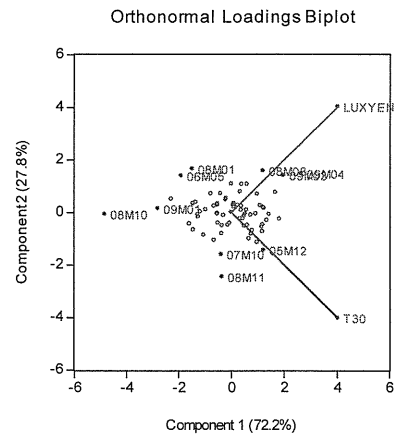
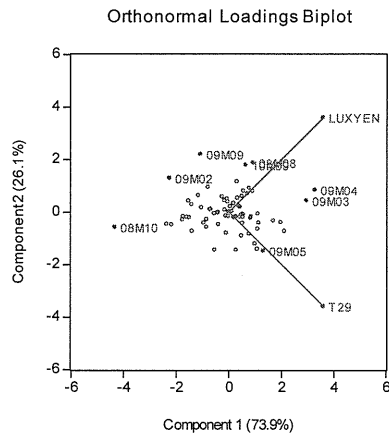
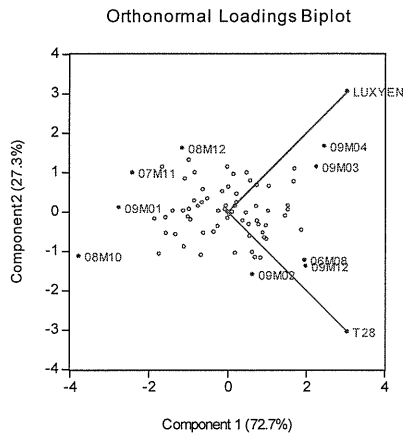
Appendix E (continued)

Orthonormal biplots of Apparel & Luxury Goods Index (yen base) versus TOPIX and its 33 sectors

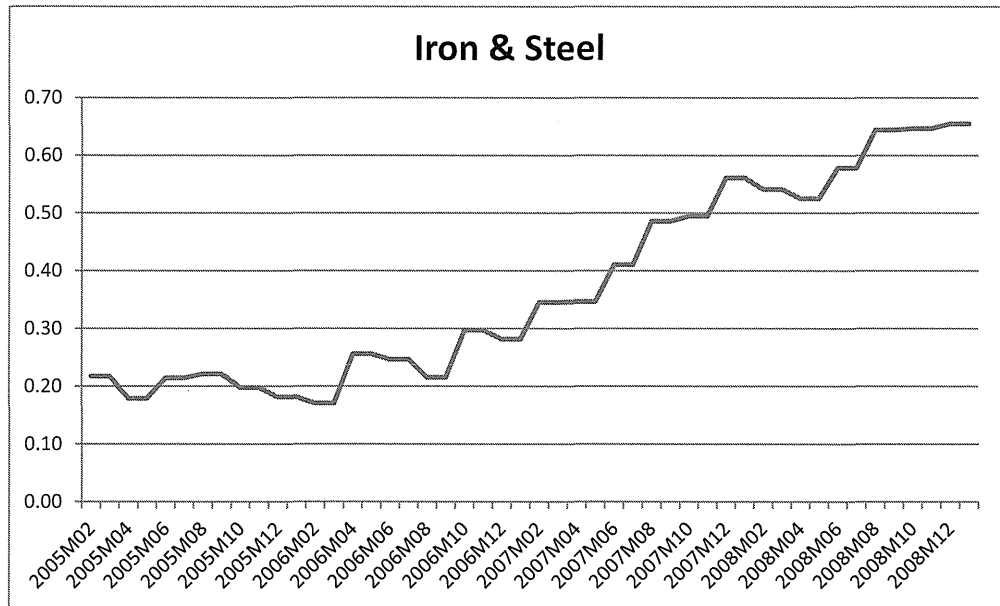


Appendix E (continued)

Orthonormal biplots of Apparel & Luxury Goods Index (yen base) versus TOPIX and its 33 sectors



Appendix F Averaged R-square of Rolling style regression estimates



高頻度取引が価格発見過程に与える影響

竹原 均

摘要

本研究においては、2010年1月4日に稼働した東京証券取引所の新取引システムであるアローヘッドにより、価格発見過程に変化が起きたかどうかを検証する。最初にティックデータを使用して、流動性の指標として一般的に用いられる、売買回転率、ビッド・アスク・スプレッド、イフェクティブ・スプレッド等の変化を確認した後に、Lesmond, Ogden and Trzcinka (1999)により提案された制限従属変数モデルを使用して限界取引コストを推定する。さらにEasley, Hvidkjaer and O'Hara (2002)のPIN, Duarte and Young (2009)のAdjusted PIN, PSOSを計算し、これらの変数を用いて投資家のオーダーフローからの私的情報の学習状況の変化を確認する。分析の結果、アローヘッドの稼働後、特に大型株を中心として流動性が上昇し、同時にスプレッドが縮小したものの、小型株に与えた影響はそれほど大きくなかったことが明らかとなった。また株式投資家間の情報の非対称性の指標とされるPIN,あるいはAdjusted PINは、アローヘッド稼働後に上昇しており、このことは取引の高速化に伴い、オーダーフローを介して私的情報が伝達されにくくなったことを示唆する。

1. アローヘッドの稼働と価格発見過程の変化

2010年1月4日に稼働した東京証券取引所の新株式売買システム「アローヘッド」(arrowhead)は、それ以前の取引システムと比較して、より高速な取引を可能とした。同システムを開発し富士通によれば、arrowheadでは5ミリ秒の注文応答時間、3ミリ秒の情報配スピードを実現しており、売買・市場情報の両面でミリ秒レベルのスピードを実現したことにより、株式売買の流動性が大幅に向上したことは間違いない。

竹原(2012)が議論しているように、流動性の向上に伴い、限界的な取引コストが低下しているとすれば、そのことは株式市場における情報効率性を高める可能性を持つ。すなわち、これまでは取引コストが障害となり、市場での取引と価格形成に使用されることなく放置されていたアルファの源泉を、トレーダーがより早い段階で利用することにより、株価がより速く均衡価格へと近づき、かつ最終的な均衡価格からの許容乖離幅が従来よりも縮小する可能性を持つ。

一方、arrowhead、ならびに東京証券取引所によるコロケーション・サービスの提供は、ヘッジファンドなど一部の投資家の高頻度取引(High Frequency Trades, HFT)を加速し、これによりザラバにおける注文電文、約定電文の状況から、目視により他の投資家の持つ私的情報を類推することが困難になったとも言われる。その結果として、HFTへの対抗手段を持たない中小規模の証券会社は主たる収益源を失うことにつながった。つまりarrowheadの稼働により市場効率性が高まったことにより、従来は残されていた取引コストにより説明可能なアルファが縮小し、この部分を収益源としていた投資家がダメージを受けたのである。

しかし竹原(2012)での分析期間は1980年1月～2010年12月であり、arrowhead稼働後1年間のデータが含まれているのみであり、arrowhead稼働が価格発見過程に与えた影響については、ほとんど言及していない。本研究では、2012年6月までのデータを使用することにより、arrowheadの稼働の前後での価格発見過程の変化を、より詳細に確認することを試みる。ここでの分析の主眼は、限界的取引コストが低下したのか、そして投資家間の情報の非対称性が縮小したのかの2点にある。

取引コストについては、先行研究の多くで利用されているbid-ask spread, effective spreadではなく、Lesmond, Ogden and Trzcinka (1999)で提案された制限従属変数モデル(Limited Dependent Variable Model, LDV model)を用いて推定された限界取引コストの時間変化を本研究では確認する。Lesmond, Ogden and Trzcinka (1999)は、もっとも低いコストで取引が可能となる限界投資家(marginal investor)であっても許容せざるを得ない取引コストを限界取引コストとしている。一方、Ogden (2009)は、LDVモデルのもとでの取引コストとは、マーケッ

ト・インパクト、手数料等を考慮した実質的な裁定コストであるとしている。いずれにしても、アルファの源泉となる投資情報が存在していたとしても、それが限界取引コスト未満の場合には、投資戦略を実行したとしても、獲得可能なアルファよりも取引コストが上回るため投資戦略の実行は経済的な利益をもたらさない。結果として投資情報が価値を持つとしても、限界投資家であってさえも、それを利用することが出来ないことから放置される。このため見掛け上のアノマリーが存在するものの、それは取引コストにより説明可能であり、実現可能な裁定機会ではない。

次に投資家間の情報の非対称性については、Easley, Hvidkjaer and O'Hara (2002)で提案された Probability of Private Information Based Trades, いわゆる 'PIN variable' を尺度として、arrowhead 稼働前後での変化を確認する。ただし、Easley, Hvidkjaer and O'Hara (2002)の PIN に関しては、Duarte and Young (2009)が批判しているように、本来は区別されるべきである個別銘柄の流動性と情報の非対称性を同時に 1 変数で測定した結果であるとも考えられる。このため Duarte and Young (2009)の拡張モデルを使用して、情報の非対称性の尺度とされる 'Adjusted PIN', そして会計情報に代表される公的情報による流動性の変化の尺度である Probability of Symmetric Order-flow Shock (PSOS)についても、これらを推定し、その時系列での変化を検証することとする。

論文は以下のように構成される。まず2節, 3節において, LDV モデルのもとでの限界取引コスト, それに PIN, Adjusted PIN, PSOS variables の推定の概要について説明する。続く 4, 5 節では, 取引コストと情報の非対称性の arrowhead 稼働の前後での変化について確認する。最後に 6 章において結論を述べるとともに, 年金資産運用へのインプリケーション, 特に株式からの直接的なキャッシュアウトを行う際の留意点について言及する。

2. 限界取引コストの推定 - LDV モデル -

Lesmond, Ogden and Trzcinka (1999)は, 前節でも議論したように保有している投資情報の価値(すなわち期待アルファ)と取引コストの差が最大となる投資家を, 限界投資家(marginal investor)と定義している。仮に投資家の保有する情報が同一であれば, 限界投資家とは, 最も低いコストで取引可能な投資家であり, その意味で情報の投資戦略における利用に関して最優位な立場にある。

ここで限界投資家が保有する証券の売却時期待取引コスト率を α_{1j} , 購入時期待取引コスト率を α_{2j} とする。投資情報の価値より限界投資家の取引コストが高い場合には, すべての投資家は投資戦略の実行により経済的な利益を獲得することが期待されないため, 株式市

場において約定は成立せず，約定価格ベースで計測された株式リターンはゼロとなる。

ここで第 j 銘柄の真のリターンを $r_{j,t,d}^*$ ，ベンチマーク・リターンを $r_{M,t,d}$ ，無危険利子率を $r_{f,t,d}$ とし，以下の資本資産価格評価モデル(Capital Asset Pricing Model, CAPM)が成立しているとする。

$$r_{j,t,d}^* - r_{f,t,d} = \beta_j (r_{M,t,d} - r_{f,t,d}) + \varepsilon_{j,t,d} \quad (1)$$

(1)式左辺の個別銘柄リターンは，情報価値がコストを超えた場合についてのみ非ゼロとなるので，これは limited dependent variable (LDV, 制限従属変数) として扱う必要がある。このアイデアに基づいて，Lesmond, Ogden and Trzcinka (1999)は，限界投資家の取引コスト推定のための LDV model を提案した。

被説明変数 $r_{j,t,d}^*$ と説明変数であるベンチマーク・リターン $r_{M,t,d}$ の関係を感覚的に示したのが図 1 である。投資戦略の実行可能性(収益性)の評価から，第 j 銘柄のリターンは，限界投資家の売却時期待取引コスト α_{1j} と購入時期待取引コスト α_{2j} の間(Region 0, 区間 $[\alpha_{1j}, \alpha_{2j}]$) でゼロとなり，そしてその外側(Region 1, Region 2)では傾き β_j の直線となるため， α_{1j} , α_{2j} の 2 点を節点とする区分線形関数となる。その一方で実現リターンは情報価値を部分的に開示するので区分線形関数に沿った S 字曲線となる。

この時に，Lesmond, Ogden and Trzcinka (1999, p.1120-1123)で示されたように，対数尤度

$$\begin{aligned} \ln L = & \sum_1 \ln \frac{1}{\sqrt{2\pi}\sigma} - \sum_1 \frac{1}{2\sigma_j^2} (R_{jt} + \alpha_{1j} - \beta_j R_{Mt})^2 \\ & + \sum_2 \ln \frac{1}{\sqrt{2\pi}\sigma} - \sum_2 \frac{1}{2\sigma_j^2} (R_{jt} + \alpha_{2j} - \beta_j R_{Mt})^2 \\ & + \sum_0 \ln(N((\alpha_{2j} - \beta_j \cdot R_{Mt})/\sigma_j) - N((\alpha_{1j} - \beta_j \cdot R_{Mt})/\sigma_j)) \end{aligned} \quad (2)$$

を数値的に最大化することによって，個別銘柄とベンチマークの日次リターンを入力として，売却時取引コスト α_{1j} ，および購入時取引コスト α_{2j} ，ベータ β_j ，ボラティリティー σ_j を最尤法により同時に推定することが可能である。この時に $(\alpha_{2j} - \alpha_{1j})$ がリバランス時の往復売買に要する限界取引コストとなる。

対数尤度(2)のもとでの最尤推定において，パラメータ数が 4 である。したがって月次推定

もサンプル数(月間取引日数)が 20 程度であることから可能ではある。しかし経験的に月次での推定値には、時間推移を確認した場合のぶれが大きいことから、本研究では1年間の移動平均による推定、すなわち過去1年間の日次リターンを入力とした推定と、月次での推定の両方を実施し、両者を比較することによって結果の頑健性を確認することを試みている。

3. 高頻度取引が情報の非対称性に与える影響 — PIN Variable —

流動性と情報の非対称性の間に正の相関が存在するとしても、両者は同一ではない。しかしながら、これまで実証分析において両者は多くの場合に混同されてきた。例えばビッド・アスク・スプレッド、あるいはイフェクティブ・スプレッドが大きな銘柄の取引コストが高いとしても、情報の非対称性の度合いが高いとは限らない。投資家の直面する情報環境が同一であり、情報の非対称性の度合いが同一であるとしても、小型株、あるいは浮動株比率が低い銘柄の取引コストは相対的に高いはずである。流動性と情報の非対称性が正の相関を持ちながらも、異なるものであるとすれば、情報の非対称性を適切に測定する尺度が実証分析において必要とされるのである。

このような情報の非対称性の尺度の候補の一つとして、実証研究において頻繁に使用されるのが、Easley, Hvidkjaer and O'Hara(2002)による probability of private information based trades, いわゆる PIN variable である。

Easley, Hvidkjaer and O'Hara(2002)のモデルでの株式市場には公的情報(public information)のみを受け取る非情報投資家 (uninformed investor) と公的情報に加えて私的情報(private information)を受け取る情報投資家の2種類が存在し、非情報投資家、情報投資家の売買注文は、ともにポワソン過程に従うものとされている。図2はモデルの基本構造を示したものであるが、まず取引日には私的情報が市場にもたらされる日と、もたらされない日がある。ここで市場に私的情報がもたらされる確率を α とする。(したがって市場に私的情報がもたらされない確率は $(1-\alpha)$ である。) 次に、私的情報が投資家にとって bad news である確率を δ , good news である確率を $1-\delta$ とする。市場に私的情報がもたらされない取引日(確率 $1-\alpha$)においては、非情報投資家だけが取引を行うものとして、その買い注文の到着率を ε_b , 売り注文の到着率を ε_s とする。一方、私的情報が市場にもたらされた取引日においては、非情報投資家に加えて、情報投資家による売買注文も到着する。しかし私的情報の内容に対応して、情報投資家は good news (signal high)の場合には株価上昇を予想して買い注文を、bad news (signal low)の場合には株価下落の予想のもとに売り注文を出す。このため情報投資家による追加的な約定到着率を μ とすれば、確率 $\alpha(1-\delta)$ で買い注文の到着率が ε_b から $\mu+\varepsilon_b$ に、

確率 $\alpha\delta$ で売り注文の到着率が ε_s から $\mu+\varepsilon_s$ に増加する。

図2から理解されるように, Easley, Hvidkjaer and O'Hara (2002)のモデルには, (1) 私的情報 (ただし bad news)が市場にもたらされる, (2) 私的情報(good news)がもたらされる, (3) 公的情報のみが市場にもたらされるという 3 状態が存在し, モデル内の推定対象となるパラメータは 5 種類($\alpha, \delta, \varepsilon_b, \varepsilon_s, \mu$)である。

このため, 以下の(3)式で与えられる尤度関数でも, 状態(1)~(3)に対応して 3 種類の項が存在する。(ただし, ここで B, S は買い主導, 売り主導約定件数。¹⁾)

$$\begin{aligned}
 L(\theta | B, S) = & \alpha\delta e^{-\varepsilon_b} \frac{\varepsilon_b^B}{B!} e^{-(\mu+\varepsilon_s)} \frac{(\mu+\varepsilon_s)^S}{S!} \\
 & + \alpha(1-\delta) e^{-(\mu+\varepsilon_b)} \frac{(\mu+\varepsilon_b)^B}{B!} e^{-\varepsilon_s} \frac{\varepsilon_s^S}{S!} \\
 & + (1-\alpha) e^{-\varepsilon_b} \frac{\varepsilon_b^B}{B!} e^{-\varepsilon_s} \frac{\varepsilon_s^S}{S!}
 \end{aligned} \tag{3}$$

ここで過去 T 日間の買い約定件数 B_t , 売り約定件数 S_t を入力として, 尤度関数の積を最大化する。つまりパラメータ $\eta=(\alpha, \delta, \varepsilon_b, \varepsilon_s, \mu)$ として,

$$\text{Maximize } \prod_{t=1}^T L(\eta | B_t, S_t) \tag{4}$$

を解くことにより, パラメータを最尤推定する。その推定結果を用いて, Easley, Hvidkjaer and O'Hara (2002)は, probability of private information based trades, いわゆる PIN variable を以下の(5)式により定義した。

$$PIN = \frac{\alpha\mu}{\alpha\mu + \varepsilon_b + \varepsilon_s} \tag{5}$$

(5)式の分母は, 非情報投資家と情報投資家の両方による約定電文の到着率であり, 分子は情報投資家からの約定の到着率である。モデルの設定上, 私的情報を有する情報投資家が取引を行うのは, 私的情報が市場にもたらされた場合に限られるので, 市場にもたらされる私的情報量が多い場合, つまりは投資家間の情報の非対称性の度合いが大きい場合に, PIN は大きな値を取る。

¹⁾ 約定電文の売り主導, 買い主導への分類と, 個別銘柄ティックデータの前処理, 買い主導・売り主導約定電文数日次データの計算については, 太田・宇野・竹原 (2011), 第8章を参照のこと。

この(5)式で与えられた PIN は、その後多くのファイナンス、会計分野の実証分析において情報の非対称性の尺度として用いられた。しかし同時に PIN が情報の非対称性の尺度であるのかディベートの対象となってきた。PIN に関する批判的な立場の研究の中で、Easley, Hvidkjaer and O'Hara (2002)によるモデルの理論的欠点を指摘し、情報の非対称性と流動性を明示的に区別したモデルを提案したのが Duarte and Young (2009)である。

Easley, Hvidkjaer and O'Hara (2002)では、私的情報のオーダーフローへの影響は、私的情報の内容に依存して買い、あるいは売りの一方のみという非対称(asymmetry)なものであった。このような私的情報が与える asymmetric order-flow shocks に対して、Duarte and Young (2009)は売り注文と買い注文の両方に影響を与えるという意味で対称(symmetric)な、公的ニュース(public news)によるオーダーフローの変化、すなわち symmetric order-flow shocks (SOS) を考慮してモデルの拡張を行った。簡単に言えば、企業が業績予想を修正し、それをアナウンスしたとすれば、それは公的情報であるが約定電文数を増加させる。そうした公的情報(あるいはイベント)が与える影響、つまりは情報の非対称性とは関係のない流動性の変化をモデルに取り込むことを試みているのである。

ここである取引日に symmetric order-flow shocks が株式市場において生起する確率を θ とする。公的な情報に基づく SOS は、投資家はその情報内容をどのように解釈するかによって、売り注文、買い注文の両方を増加させるため、その買い注文の到着率の上昇分を Δ_B 、売り注文の到着率の上昇分を Δ_S とする。さらに Easley, Hvidkjaer and O'Hara (2002)では、私的情報がオーダーフローに与える影響は売り注文、買い注文に共通して μ であるとしていたが、これを私的情報が good news であることによる買い注文到着率の上昇分 μ_B 、bad news であることによる売り注文到着率の上昇分を μ_S と区別する。こうした SOS によるオーダーフロー変化を組み込むことにより、Duarte and Young (2009)の拡張モデルの基本構造を表した図3を得る。この図からわかるように、このモデルでは SOS が起きた場合に限り、買い注文、売り注文の到着率がそれぞれ Δ_B, Δ_S だけ上昇するため、モデル内の状態数が2倍に増えているものの、それ以外は Easley, Hvidkjaer and O'Hara (2002)のモデルと μ_B, μ_S を区別した点を除いてまったく同一である。

ここで推定対象パラメータの集合を、 $\eta = (\alpha, \delta, \theta, \varepsilon_b, \varepsilon_s, \mu_b, \mu_s, \Delta_b, \Delta_s)$ とするとき、Duarte and Young (2009)のモデルにおける尤度関数は以下の(6)式で与えられる。