Online Appendix

A.1 Example of Disaggregated SG&A Information

This appendix displays two examples of disaggregated SG&A disclosure of firms that provide and that do not provide disaggregated SG&A information. The former is FANKL CORPORATION and, the latter is MEGMILK SNOW BRAND Co., Ltd.

FANKL CORPORATION

FANKL CORPORATION is a Japanese company that has developed additive-free cosmetics, supplements, sprouted brown rice and kale juice. Figure A-1 displays the corporation's *Kessantanshin* for 2023. The items surrounded by the dotted lines are disaggregated SG&A information. The company believes that reliability is an important factor when purchasing products, especially supplements, and has therefore focused on creating a good brand image. Accordingly, it spent 14,868 million yen as advertising expenses in 2023, accounting for approximately 24% of total SG&A expenses. As these marketing activities will benefit the company in the long term, disclosing this information is considered important to distinguish it from current expenditures included in SG&A expenses.

Figure A-1 FANKL Corporation Example

(2) Consolidated Statement of Income and Consolidated Statement of Comprehensive Income

Consolidated Statement of Income

(Millions of yen, rounded down)

	Fiscal year ended March 31, 2023	Fiscal year ended March 31, 2022
Net sales	103,595	103,992
Cost of sales	34,382	34,876
Gross profit	69,213	69,116
Selling, general and administrative expenses		
Sales promotion expenses	6,887	7,645
Packing and transport expenses	5,478	5,448
Advertising expenses	14,868	12,577
Sales commission fee	3,111	2,933
Communications expenses	1,711	1,688
Directors' remuneration	415	426
Salaries and bonuses	11,016	11,075
Provision for accrued bonuses	1,027	1,006
Provision for share-based remuneration	156	108
Retirement benefit expenses	525	505
Compulsory welfare expenses	1,750	1,763
Welfare expenses	276	286
Depreciation	2,612	2,611
Research and development expenses	1,225	1,257
Rent expenses	1,172	1,216
Provisions for allowance for bad debt	28	43
Other	9,105	8,749
Total selling, general and administrative expenses	61,370	59,346
Operating income	7,843	9,771

MEGMILK SNOW BRAND Co., Ltd

MEGMILK SNOW BRAND Co., Ltd is a Japanese company that produces and sells milk products, including butter, margarine, cheese, milk and milk beverages, yogurt, and other products. Figure A-2 displays the company's disaggregated SG&A information in *Kessantanshin* for 2017. The largest amount is the sales promotion cost followed by transport and storage fees, which amounted to 72,005 million yen in 2017 and accounts for about 59% of the total SG&A expenses (which were 123,017 million yen in 2017). Compared with advertising targeting the general public and intending to raise brand value, sales promotion costs target specific customers aiming to increase sales for the current period, rather than for the long term. Moreover, transport and storage fees are more likely to be expenditures for the current revenue. As most SG&A expenses matched with current revenue, the demand for disaggregated SG&A information is predicted to be low. Consistent with this prediction, the company disclosed disaggregated SG&A information in *Kessantanshin* until 2017, but not thereafter.

Figure A-2 MEGMILK SNOW BRAND Co., Ltd Example

Note 2: Major expense items and amounts for selling expenses are as follows:

	Year ended March 31, 2016	Year ended March 31, 2017
Salaries	9,842 million yen	9,730 million yen
Bonus allowances	1,335	1,365
Provision for bonuses	1,623	1,826
Retirement benefit expenses	803	720
Transport & storage fees	33,256	34,167
Advertising costs	4,588	5,397
Sales promotion costs	37,246	37,838
Provision for doubtful accounts	(77)	83
Depreciation and amortization costs	644	716
Other costs	13,950	14,057
Total	103,214	105,905

Note 3: Major expense items and amounts for general administrative expenses are as follows:

	Year ended March 31, 2016	Year ended March 31, 2017
Salaries	4,501 million yen	4,425 million yen
Bonus allowances	659	615
Provision for bonuses	887	987
Retirement benefit expenses	305	272
Depreciation and amortization costs	1,526	1,508
Rents	369	362
Other costs	8,383	8,939
Total	16,633	17,112

A.2 Estimation of SGAFV

This paper estimated *SGAFV* by firm-year in a two-step process following Banker et al. (2019). The first step estimates the optimal lag structure of SG&A expenses by estimating how many later periods the current SG&A expenses are linked to profit in each industry. In the second step, it estimates the firm-year future value of SG&A expenses using the lag structure obtained in the first step as a set of given values.

First, the study estimates the optimal lag structure of SG&A expenses in each industry sector. For 2000-2023, it estimates equation (A-1) including year dummies for each industry sector (*Nikkei medium classification*):

$$\frac{OI_{i,t}}{A_{i,t}} = \alpha_0 + \alpha_1 \frac{1}{A_{i,t-1}} + \sum_{k=0}^{n} \alpha_{2,k} \frac{SGA_{i,t-k}}{A_{i,t-k}} + \varepsilon_{i,t}, \tag{A-1}$$

where OI is the sum of operating income and SG&A expenses, A is the total assets at the end of the period, and SGA is the SG&A expenses.

This study uses industry average of SG&A expenses as *SG&A* to address simultaneity issues (Banker et al. 2019). For example, suppose a shock increases demand for a firm's product. At this time, the earnings (the dependent variable) increase owing to this rise. Simultaneously, the independent variables may increase in response to the increase in demand. For each year and industry (*Nikkei medium classification*), it regress SG&A expenses on the average SG&A expenses of other firms classified in the same industry (*Nikkei small classification*) to which the firm belongs:

$$\frac{SGA_{i,t}}{A_{i,t}} = \alpha_0 + b \frac{SGA\_NDI_{i,t}}{A\_IND_{i,t}} + \varepsilon_{i,t}, \tag{A-2}$$

where $A\_IND_{i,t}$ represents the average total assets other than firm i in the industry (*Nikkei small classification*) to which firm i belongs in year t and $SGA\_IND_{i,t}$ represents the average value of SG&A expenses other than firm i in the industry (*Nikkei small classification*) to which firm i belongs in year t. From the coefficients obtained, this study predicts $SGA_{i,t}/A_{i,t}$ and use it in the estimation of equation (A-1) instead of the actual value.

To estimate the optimal lag structure, this paper estimates equation (A-1) for each industry (*Nikkei medium classification*), with eight different models ranging from k=0 to k=7. Among the eight models, it considers the model in which all coefficients of the lag in SG&A expenses are significant (10% in a one-tailed test) and positive a candidate for the optimal lag structure. Among the selected candidates, it compares the Akaike information criterion (AIC) and Bayesian information criterion (BIC) , and selects the model with the lowest value; if the AIC and BIC results are inconsistent, it selects the model with the highest adjusted coefficient of determination.

In the second step, it estimates equation (A-1) for each firm-year using the lag structure model

selected in the first step to estimate the future value of SG&A expenses for each firm-year. As in the first step, this study uses instrument variables instead of firms' actual SG&A. Using the most recent data since 2000, it estimates (A-1) for each firm year. For example, when estimating the future value of SG&A expenses for firm i in 2011, it uses data from 2000 to 2011 for firm i to perform the estimation. This study excludes those for which 10 or more observations are unavailable from the sample. Thus, data on future values have been available since 2009. Assuming a discount rate of 10%, it defines the future value of SG&A expenses as $\frac{\sum_{k=1}^{n} \frac{\alpha_{2,k}}{(1.1)^k} * SGA_{i,t}}{A_{i,t}}$, which is calculated from the obtained coefficients. The future value of SG&A expenses indicates how much of the SG&A expenses spent in the current year generate future earnings.

A.3 Estimation of *ABDE_all* and *ABDE_five*

To measure real earnings management through SG&A expenses, this paper estimate equation (A-3) for each industry-year, following Roychowdhury (2006):

$$\frac{DE_{i,t}}{A_{i,t-1}} = \alpha_0 + \alpha_1 \frac{1}{A_{i,t-1}} + \beta_1 \frac{S_{i,t-1}}{A_{i,t-1}} + \varepsilon_{i,t}, \tag{A-3}$$

where DE and S represent the discretionary expenditures and sales, respectively. This study calculates DE in two ways: using total SG&A expenses (*ABDE_all*), following Roychowdhury (2006), and sum of five SG&A expenses: (1) R&D, (2) advertising, (3) sales promotion and other selling expenses, (4) directors' compensation and bonuses, and (5) personnel expenses (*ABDE_five*), following Yamaguchi (2021). This study estimates equation (A-3) for each industry year and uses the obtained coefficients to calculate normal discretionary expenditure for each firm-year. Next, it calculate abnormal discretionary expenditures by subtracting normal discretionary expenditures from actual expenditures for each firm year. To measure the average extent of earnings management through SG&A expenses, it calculates the average abnormal discretionary expenditure over the past 5 years by firm-year. Moreover, for easy comprehension of the results, it multiplies the average value of abnormal discretionary expenditures by -1 so that larger values indicate greater real earnings management.

A.4 Fixed Effect Model Analysis

As mentioned in the Section 5, it is possible for certain firms with specific characteristics, such as a high level of corporate governance, to be more likely to provide disaggregated SG&A information. For example, Boone and White (2015) find a positive relationship between institutional ownership and management disclosure, interpreting that institutional investors need more public information to minimize transaction costs and managers respond to their demands. Therefore, firms

with higher levels of corporate governance are more likely to provide disaggregated SG&A information, because corporate governance factors encourage firms to disclose more public information.

This possibility raises endogeneity concern. First, the positive relationship between the disclosure of disaggregated information and the future asset value of SG&A expenses (Hypothesis 1) may stem from corporate governance factors. Chen, et al. (2012) find that managerial empire building, the tendency to increase (not to decrease) SG&A expense compare with the increase (decrease) of sale for pursuit of managerial personal utility, is mitigated by corporate governance. Therefore, firms with higher levels of corporate governance may be more likely to generate future asset value of SG&A expenses, because the SG&A expenditure for the managerial personal utility is restrained due to the high level of corporate governance. This prediction raises concern that the positive relationship between the disclosure of disaggregated information and the future value of SG&A simply means that firms with a high level of corporate governance are more likely to disclosure information and generate the future value of SG&A.

Second, the negative relationship between disclosure of disaggregated information and real earning management (Hypothesis 2) may stem from corporate governance factors. Numerous studies have shown that corporate governance restrains real earnings management. For example, Roychowdhury (2006) finds that institutional ownership, considered a sophisticated investor, lowers the reduction of discretionary spending. In addition, using Japanese sample, Guo et al. (2015) find that foreign investors' ownership lowers the reduction of the discretionary spending because they improve corporate governance by exercising voting rights. Therefore, there is a possibility that the relationship between disclosure of disaggregated information and real earnings management is negative because firms with high corporate governance are more likely to disclose information and less likely to conduct real earnings management.

Overall, the specific characteristics of the company such as the level of corporate governance may affect the results of this study. To mitigate this endogeneity concern, this study conducts a supplemental analysis by adding firm fixed effects in the main regression analysis to control firm characteristics that remain constant over time and addresses the endogeneity issues caused by correlation between the independent variables and unobserved firm-specific characteristics. Of course, some firms' characteristics are not necessarily constant over time; however using a fixed effects model is effective because corporate culture and corporate climate towards compliance are unlikely to change quickly, and appear to influence firms' characteristics to some extent. Nevertheless, fixed effects models have the disadvantage of focusing on within-firm variation, even though the decision of disclosing or not disclosing disaggregated information of SG&A expenses in *Kessantanshin* is persistent (96% of the samples have continued to use the same disclosure policy as that used last year). Therefore, a cross-sectional analysis focusing on the differences between firms is more appropriate than a fixed-effect model analysis which allows for a clearer under-

standing of how changes within a firm impact the outcome. Despite the limitation of fixed effect model, this paper uses this model as a supplement of the main cross-sectional analysis to control firm characteristics that remain constant over time.

Table A-1 shows the results of the fixed effect model. Fixed effect models cannot include companies that disclose and do not disclose disaggregated SG&A information for all periods of the sample period. Therefore, the sample size decreased significantly compared to Table 2. Columns (1)-(5) show the results of whether firms with higher future asset values for SG&A expenses are more likely to disclose disaggregated SG&A information in *Kesssantanshin* (Hypothesis 1). Column (1) shows the results using *SGAINT* as the explained variable and columns (2)-(4) show the results of decomposing *SGAINT* into *Personal Expenses*, *Advertising Expenses*, and *R&D Expenses*. The coefficients of *SGAINT*, *Personal Expenses*, and *Advertising Expenses* are statistically significant at the 1% level and that of *R&D Expenses* is statistically significant at the 10% level consistent with the results of Table 2. Column (5) presents the result of *SGAFV*. The coefficient of *SGAFV* is not statistically significant, implying that firm characteristics that remain constant over time may affect the result of Table 2. Columns (6)-(8) show the results of testing whether managers are less likely to disclose disaggregated SG&A information in *Kessantanshin* when they routinely conduct earnings management through SG&A expenses (Hypothesis 2) and when the administrative burden for preparing financial reports is higher (Hypothesis 3). The coefficients of *ABDE_all* in column (6) and *ABDE_five* in column (7) are negative at the 10% and 1% significance level respectively and *SUBSIDIARY* in colunn (8) is negative at the 1% significance level, consistent with results of Table 2. In summary, fixed effects model analysis shows the results for all variables except *SGAFV* are similar to those in Table 2, suggesting that the results of this study are robust even when endogeneity is considered.

Table A-1 Results of firm fixed model

Dependent variable = $DISCLOSURE_{i,t}$	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
$SGAINT_{i,t}$	18.201***							
	(3.606)							
Personal Expenses$_{i,t}$		15.788***						
		(4.475)						
Advertising Expenses$_{i,t}$			42.673***					
			(9.101)					
R&D Expenses$_{i,t}$				17.783*				
				(10.595)				
$SGAFV_{i,t}$					0.973			
					(1.168)			
$ABDE\_all_{i,t}$						-5.549*		
						(3.090)		
$ABDE\_five_{i,t}$							-17.966***	
							(4.689)	
$SUBSIDIARY_{i,t}$								-25.639***
								(4.899)
Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	10,419	10,419	10,419	10,419	1,804	10,419	10,419	10,419
pseudo R-sq	0.729	0.727	0.728	0.726	0.816	0.726	0.728	0.729
Log-likelihood	-1298.745	-1305.349	-1301.608	-1310.489	-147.067	-1309.742	-1304.133	-1296.834

Note: The explained variable *DISCLOSURE* is a dummy variable that takes a value of 1 if firms disclose disaggregated SG&A information on *Kessantanshin* and 0 otherwise. Control variables include *SALES_SD*, *SGA_SD*, *RET_SD*, *LOSS*, *FE*, *AGE*, *SPEED*, *FRGN*, *%PEERS*, *MV*, *BM*, *SGA*, industry dummy, and year dummy. To avoid multicollinearity problems *SGA* is excluded from the control variables in the analysis of $SGAINT_{i,t}$ and $SGAFV_{i,t}$. Table 1 show the definitions of variables are presented in Table 1. The numbers in parentheses are the clustered standard errors. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

A.5 Tables

Table A-2 Sample selection

Firms with a fiscal year ending in March for 2009-2023 and those submitting *Kessantanshin* within 50 days of closing business.	32,263

Sample of the analysis without *SGAFV*		Sample of the analysis with *SGAFV*	
Exclude observations for which the data necessary to estimate *ABDE* cannot be obtained	27,158	Exclude industries for which the data necessary to estimate *SGAFV* cannot be obtained	31,123
Exclude observations for which control variables are unavailable	26,744	Exclude observations for which the data necessary to estimate *SGAFV* cannot be obtained	22,682
Exclude ±0.5% of continuous variables	**24,357**	Exclude industries with an estimated lag coefficient of 0	15,027
		Exclude observations with negative *SGAFV*	6,064
		Exclude observations for which control variables are unavailable	6,011
		Exclude ±0.5% of continuous variables	**5,370**

Table A-3 Summary statistics

	N	Mean	SD	p25	p50	p75
$DISCLOSURE_{i,t}$	24,357	0.577	0.494	0.000	1.000	1.000
$SGAFV_{i,t}$	5,370	0.207	0.362	0.033	0.092	0.231
$SGAINT_{i,t}$	24,357	0.086	0.068	0.040	0.066	0.108
$ABDE\_all_{i,t}$	24,357	0.008	0.111	−0.030	0.016	0.061
$ABDE\_five_{i,t}$	24,357	0.004	0.060	−0.016	0.009	0.033
$SUBSIDIARY_{i,t}$	24,357	0.023	0.027	0.007	0.015	0.029
$SALES\_SD_{i,t}$	24,357	0.081	0.067	0.038	0.062	0.101
$SGA\_SD_{i,t}$	24,357	0.015	0.018	0.005	0.009	0.017
$RET\_SD_{i,t}$	24,357	0.023	0.010	0.016	0.021	0.027
$LOSS_{i,t}$	24,357	0.116	0.320	0.000	0.000	0.000
$FE_{i,t}$	24,357	0.022	0.039	0.004	0.010	0.023
$AGE_{i,t}$	24,357	3.993	0.508	3.807	4.143	4.304
$SPEED_{i,t}$	24,357	−3.674	0.190	−3.784	−3.738	−3.664
$FRGN_{i,t}$	24,357	0.072	0.091	0.005	0.029	0.110
$\%PEERS_{i,t}$	24,357	0.568	0.187	0.415	0.560	0.724
$MV_{i,t}$	24,357	9.785	1.555	8.634	9.611	10.801
$BM_{i,t}$	24,357	1.340	0.780	0.771	1.195	1.760
$SGA_{i,t}$	24,357	0.202	0.148	0.106	0.160	0.245

Table A-4 Correlation matrix

		[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]	[16]	[17]
[1]	$DISCLOSURE_{i,t}$	1.00																
[2]	$SGAINT_{i,t}$	0.13	1.00															
[3]	$ABDE\_all_{i,t}$	−0.01	−0.58	1.00														
[4]	$ABDE\_five_{i,t}$	−0.08	−0.64	0.89	1.00													
[5]	$SUBSIDIARY_{i,t}$	−0.03	0.06	−0.08	−0.06	1.00												
[6]	$SALES\_SD_{i,t}$	0.03	0.13	0.04	0.05	0.02	1.00											
[7]	$SGA\_SD_{i,t}$	0.06	0.53	−0.37	−0.31	0.19	0.45	1.00										
[8]	$RET\_SD_{i,t}$	0.08	0.03	0.04	0.04	0.18	0.19	0.20	1.00									
[9]	$LOSS_{i,t}$	0.03	0.06	−0.03	−0.03	0.13	0.10	0.12	0.21	1.00								
[10]	$FE_{i,t}$	0.03	0.00	0.01	0.02	0.08	0.10	0.05	0.17	0.41	1.00							
[11]	$AGE_{i,t}$	−0.09	−0.27	0.01	0.00	−0.14	−0.23	−0.36	−0.18	−0.03	−0.00	1.00						
[12]	$SPEED_{i,t}$	−0.03	−0.01	0.02	0.02	−0.12	−0.01	−0.03	−0.05	−0.06	−0.10	−0.04	1.00					
[13]	$FRGN_{i,t}$	−0.08	−0.03	−0.03	−0.03	−0.09	−0.07	−0.05	−0.09	−0.11	−0.15	−0.02	0.11	1.00				
[14]	$\%PEERS_{i,t}$	0.37	0.18	0.00	0.00	−0.06	0.17	0.16	0.21	0.10	0.09	−0.19	−0.01	−0.19	1.00			
[15]	$MV_{i,t}$	−0.12	−0.09	−0.09	−0.08	−0.24	−0.16	−0.14	−0.18	−0.18	−0.26	0.12	0.19	0.62	−0.17	1.00		
[16]	$BM_{i,t}$	0.05	−0.20	0.09	0.08	−0.03	−0.07	−0.23	−0.12	0.13	0.24	0.18	−0.12	−0.26	0.12	−0.43	1.00	
[17]	$SGA_{i,t}$	0.09	0.91	−0.63	−0.57	0.08	0.17	0.60	0.02	0.07	0.00	−0.29	−0.03	−0.06	0.19	−0.10	−0.22	1.00