Harbingers of Failure: Online Appendix

Eric Anderson

Northwestern University

Kellogg School of Management

Song Lin

MIT

Sloan School of Management

Duncan Simester

MIT

Sloan School of Management

Catherine Tucker

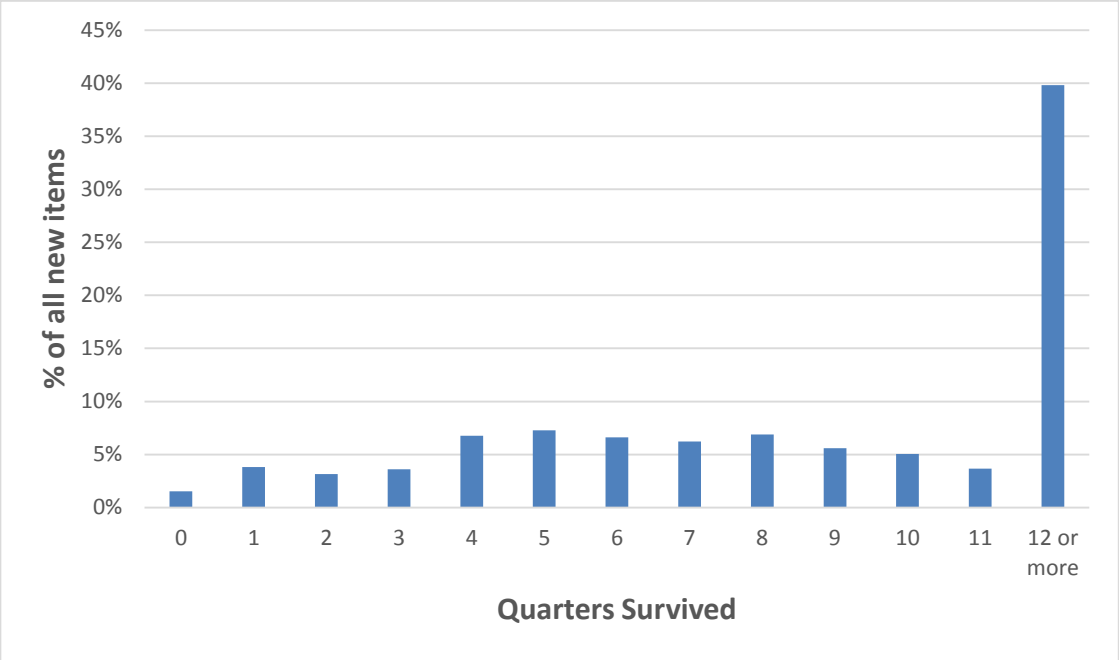
MIT

Sloan School of Management

February 2015

List of Tables in the Web Appendix

Figure WA1: Histogram of New Product Survival Durations



This figure is a histogram reporting how many quarters each new product survived. The figure is constructed using the sample of 8,809 new items in the customer transaction data.

Table WA1: Variable Definitions: Product Covariates

Pricing and Promotion	
Price Paid ($)	Average price paid. In our multivariate analysis we use the log transformation of this price.
Profit Margin (%)	Average profit margin, calculated as the average of: (regular price – unit cost)/regular price.
Discount Received (%)	Average promotion depth, calculated as the average of: (regular price – price paid)/regular price.
Discount Frequency (%)	Percentage of purchases bought on promotion.
Competitive Environment and Sales Volume	
Herfindahl Index	Sum of the squared market shares (unit sales) for each product.
Category Sales	Total unit sales in the category. In our multivariate analysis we use the log transformation of this unit sales measure.
Vendor Sales	Total unit sales of products with this vendor. In our multivariate analysis we use the log transformation of this unit sales measure.
Product Characteristics	
Private Label	A binary indicator revealing whether the product has the store's private label brand (1=Yes; 0=No).
Repeat Measures	
Customers with *n* Repeats	Number of customers who purchased the new product at least *n* times.

Table WA2: Summary Statistics: Product Covariates

	Mean	Standard Deviation
Pricing and Promotion		
Price Paid ($)	6.929	8.341
Profit Margin (%)	47.281	15.087
Discount Received (%)	11.307	17.785
Discount Frequency (%)	21.728	28.412
Competitive Environment and Sales Volume		
Herfindahl Index	0.065	0.137
Category Sales (10,000s)	0.673	0.671
Vendor Sales (10,000s)	0.620	0.910
Product Characteristics		
Private Label	0.177	0.382
Repeat Measures		
Num. Customers with 1 repeat	1.233	3.654
Num. Customers with 2 repeats	0.172	0.778
Num. Customers with 3 or more repeats	0.086	0.476

These product and category measures are calculated using the weekly store purchases. The data includes the 2,953 new products included in the regression analysis. The pricing, promotion, and repeat measures are constructed using purchases in the 15-week initial valuation period. There are 288 products with no sales during this period; they are excluded from the calculation. Competitive environment and sales volume measures are constructed for all the 2,953 products using all purchases in the classification period (November 2003 to July 2004).

Table WA3: 2-year Survival Window

	Model 1	Model 2
Total Sales	0.0011*	
	(0.0005)	
Group 1 Sales		0.0012
		(0.0031)
Group 2 Sales		-0.0110**
		(0.0037)
Group 3 Sales		-0.0180**
		(0.0069)
Group 4 Sales		-0.0158
		(0.0122)
Sales to Other Customers		0.0133**
		(0.0025)
Log Likelihood	-1,988	-1,951
Likelihood Ratio Test, Chi^2 (df=4)		75.35**
Area under ROC curve	0.5988	0.5981

The table reports average marginal effects from models where the dependent variable is a binary variable indicating whether the new product succeeded (1 if succeeded, 0 if failed). The definition of product success is based on 2-year window rather than 3-year window. Robust standard errors (clustered at the category level) are reported in parentheses. The unit of analysis is a new product. The sample size is 2,953. Significantly different from zero (or significant difference between Models 1 and 2): * $p<0.05$, ** $p<0.01$.

Table WA4: Using Market Share to Define Success

	Greater than 25th Percentile		Greater than 50th Percentile		Greater than 75th Percentile	
	Model 1	**Model 2**	**Model 1**	**Model 2**	**Model 1**	**Model 2**
Total Sales	0.0011**		0.0011**		0.0007**	
	(0.0003)		(0.0003)		(0.0002%)	
Group 1 Sales		0.0084		0.0085		0.0082*
		(0.0050)		(0.0046)		(0.0033)
Group 2 Sales		0.0016		0.0005		0.0009
		(0.0053)		(0.0047)		(0.0030)
Group 3 Sales		-0.0076*		-0.0078*		-0.0050
		(0.0034)		(0.0032)		(0.0026)
Group 4 Sales		-0.0224**		-0.0192**		-0.0089
		(0.0054)		(0.0049)		(0.0046)
Sales to Other Customers		0.0116**		0.0109**		0.0047**
		(0.0022)		(0.0020)		(0.0013)
Log Likelihood	-1,975	-1,933	-1,737	-1,691	-1,144	-1,116
Likelihood Ratio Test		83.09**		91.56**		57.59**
Area under ROC curve	0.6035	0.6157	0.6035	0.6150	0.6035	0.6188

The table reports average marginal effects from models where the dependent variable is a binary variable indicating whether the new product succeeded (1 if succeeded, 0 if failed). Robust standard errors (clustered at the category level) are reported in parentheses. The unit of analysis is a new product. The sample size is 2,953. Significantly different from zero (or significant difference between Models 1 and 2): * $p<0.05$, ** $p<0.01$.

Because the share of the category is confounded by the size of the category, we translate the shares into indicator variables as a measure of success:

Greater than 25th Percentile: Equals 1 if the market share is greater than the 25th percentile in the category; and zero otherwise.

Greater than 50th Percentile: Equals 1 if the market share is greater than the 50th percentile in the category; and zero otherwise.

Greater than 75th Percentile: Equals 1 if the market share is greater than the 75th percentile in the category; and zero otherwise.

Since the initial evaluation period occurs within 15 weeks of the new product introduction, sales in this evaluation period do not affect the market share measures. This ensures that purchases that affect the independent variables do not also contribute to the dependent variable.[1]

[1] If a new product is discontinued before the 1, 2 or 3-year threshold it has zero sales and so will receive a rank of 0. Any other discontinued products are omitted and not included when calculating the share ranks.

Table WA5a: Hazard Ratios

	5-week Initial Period	10-week Initial Period	15-week Initial Period
Model 1			
Total Sales	(0.995, 1.000)	(0.995, 0.998)	(0.995, 0.998)
Constant	(0.002, 0.003)	(0.003, 0.004)	(0.005, 0.006)
Period 2	(1.649, 2.323)	(1.256, 1.720)	(1.050, 1.409)
Period 3	(2.111, 2.957)	(1.598, 2.178)	(1.319, 1.764)
Period 4	(2.080, 2.950)	(1.601, 2.209)	(1.267, 1.724)
Period 5	(1.916, 2.768)	(1.406, 1.988)	(1.031, 1.450)
Period 6	(1.777, 2.673)	(1.259, 1.930)	(0.986, 1.593)
Log Likelihood	-3,465	-3,605	-3,795
Model 2			
Group 1 Sales	(0.959, 1.008)	(0.970, 1.033)	(0.943, 0.992)
Group 2 Sales	(0.968, 1.010)	(0.942, 0.995)	(0.965, 1.008)
Group 3 Sales	(1.033, 1.070)	(1.017, 1.061)	(1.010, 1.043)
Group 4 Sales	(1.088, 1.135)	(1.053, 1.109)	(1.050, 1.091)
Sales to Other Customers	(0.967, 0.985)	(0.942, 0.966)	(0.957, 0.976)
Constant	(0.002, 0.003)	(0.003, 0.004)	(0.005, 0.006)
Period 2	(1.656, 2.333)	(1.267, 1.735)	(1.061, 1.424)
Period 3	(2.132, 2.987)	(1.631, 2.223)	(1.351, 1.807)
Period 4	(2.104, 2.984)	(1.644, 2.269)	(1.306, 1.776)
Period 5	(1.942, 2.807)	(1.451, 2.051)	(1.069, 1.504)
Period 6	(1.805, 2.715)	(1.302, 1.997)	(1.027, 1.659)
Log Likelihood	-3,447	-3,563	-3,742
Comparing Model 1 and 2			
Likelihood Ratio Test, Chi^2 (df=4)	34.93^{**}	84.21^{**}	106.29^{**}

The table reports 95% confidence intervals for hazard ratios obtained from a hazard model estimating the probability a new product fails. Failure is defined as the last week in which the item has sales. The model is estimated using a common sample of 2,953 items in the prediction set. For each product the sequence starts in week 16 and continues until failure or week 156 (3-years after introduction). Significant difference between Models 1 and 2: * $p<0.05$, ** $p<0.01$.

We define a failure as the last week of transactions for an item; an item that has its last purchase in week 120 "fails" in week 120. We then use sales during the initial evaluation period to predict the hazard of a failure in each week of a product's life. The unit of analysis is an item x week and the data includes the complete sequence of each product's life. The sequence starts with the week after the initial evaluation period and continues until either the product fails or the product has survived for 3 years (156 weeks). This is a single-failure hazard model; once a product has been discontinued we do not observe additional sequences. This model naturally accommodates censoring of the data once

an item fails. The hazard rate is the probability of a failure conditional on the product surviving to that week and this probability is unaffected by the outcomes in subsequent weeks.

To ensure that the hazard function captures dynamics in the likelihood of failure phenomenon, we construct dummy variables identifying each six-month period in the first 3 years of a product's life:

Period 1	week 16 through week 26
Period 2	week 27 through week 52
Period 3	week 53 through week 78
Period 4	week 79 through week 104
Period 5	week 105 through week 130
Period 6	week 131 through week 156

We then use maximum likelihood to estimate nonparametric exponential hazard functions specified as:

Model 1: $\lambda_{it} = \exp [\alpha + \sum \beta_t \text{ Period } x_t + \beta_1 \text{Total Sales}_i]$

Model 2: $\lambda_{it} = \exp [\alpha + \sum \beta_t \text{ Period } x_t + \beta_1 \text{Group 1 Sales}_i + \beta_2 \text{Group 2 Sales}_i + \beta_3 \text{Group 3 Sales}_i + \beta_4 \text{Group 4 Sales}_i + \beta_5 \text{Sales to Other Customers}_i]$

The *Group x Sales* variables are the same variables that we used in the previous logistic model. In order to include the constant term we omit the fixed effect for period 1. The findings are summarized in the table above, where we report the 95% confidence intervals for the hazard ratios for the coefficients of interest. We estimate both models using the 5, 10, and 15 week initial evaluation periods for the 2,953 products in the prediction set.

Recall that a hazard ratio larger than one indicates that increasing values of the variable are associated with a higher hazard of failure, while a hazard ratio less than one indicates the reverse.

What makes these results particularly surprising is that they reverse the standard argument that items with higher sales are more likely to succeed. If customers' classification set purchases contained no information about the hazard of failure then the hazard ratios should be less than 1 and of similar magnitudes for all four customer groups. Recall also that regression to the mean would suggest that customers who purchased flops from the classification set would be less likely to purchase flops from the prediction set. This suggests we should observe lower hazard ratios for customers in Groups 3 and 4, and higher hazard ratios for customers in groups 1 and 2. We observe the opposite pattern.

The fit of the models (measured by the log likelihood) also reveals an interesting pattern; the shorter the initial evaluation period the better the model fit. Because we use exactly the same data in each of the models, this pattern cannot be attributed to mere sample differences. One interpretation is that adoption by harbingers is more informative if they choose to adopt the product earlier. To investigate this possibility we disaggregated the 15 week initial evaluation period (after the product was introduced) into three groupings: (i) purchases in weeks 1-5, (ii) purchases in weeks 6-10, and (iii) purchases in weeks 11-15. We then re-estimated the hazard model when counting purchases by each *FlopAffinity* group within each of these initial purchasing windows. The results are reported below.

Table WA5b: Hazard Function: Disaggregating the 15-Week Initial Evaluation Period

	Weeks 1-5	Weeks 6-10	Weeks 11-15
Model 2			
Group 1 Sales	(0.956, 1.054)	(0.981, 1.069)	(0.843, 0.929)
Group 2 Sales	(0.920, 1.006)	(0.934, 1.015)	(0.964, 1.045)
Group 3 Sales	(1.002, 1.072)	(1.002, 1.062)	(0.976, 1.034)
Group 4 Sales	(1.047, 1.139)	(1.017, 1.094)	(1.044, 1.131)
Sales to Other Customers			(0.956, 0.974)
Constant			(0.005, 0.006)
Period 2			(1.068, 1.434)
Period 3			(1.365, 1.826)
Period 4			(1.326, 1.804)
Period 5			(1.090, 1.534)
Period 6			(1.050, 1.696)
Log Likelihood			-3,720

The table reports 95% confidence intervals for hazard ratios obtained from a hazard model estimating the probability a new product fails. Failure is defined as the last week in which the item has sales. The model is estimated using a common sample of 2,953 items in the prediction set. For each product the sequence starts in week 16 and continues until failure or week 156 (3 years after introduction).

As we would expect the coefficients within each *FlopAffinity* group are similar across the three time periods. However, it does appear that purchases by customers in Group 3 are more informative in weeks 1-5 and 6-10 than in weeks 11-15. This may be sufficient to explain the improved explanatory power when we restrict the attention to a shorter initial evaluation period. We are cautious in drawing strong conclusions from this analysis, particularly as we do not see the same pattern in the main logistic model (see Table 1 and the "Varying the Initial Evaluation Period" table in the Appendix).

Table WA6: Varying the Initial Evaluation Period

	5-Weeks		10-Weeks	
	Model 1	**Model 2**	**Model 1**	**Model 2**
Total Sales	0.0034**		0.0032**	
	(0.0014)		(0.0009)	
Group 1 Sales		0.0045		-0.0046
		(0.0099)		(0.0053)
Group 2 Sales		0.0045		0.0063
		(0.0090)		(0.0065)
Group 3 Sales		0.0009		-0.0033
		(0.0065)		(0.0043)
Group 4 Sales		-0.0195*		-0.0176**
		(0.0079)		(0.0059)
Sales from Other Customers		0.0110*		0.0144**
		(0.0046)		(0.0029)
No Sales in the first 15 weeks	0.0792	0.0639	0.1542*	0.1441
	(0.0802)	(0.0800)	(0.0768)	(0.0758)
(log) Price Paid	0.0330	0.0281	0.0549**	0.0509**
	(0.0209)	(0.0210)	(0.0181)	(0.0180)
Profit Margin	0.0033	-0.0096	0.0537	0.0418
	(0.1282)	(0.1265)	(0.1248)	(0.1227)
Discount Received	-0.2956	-0.3147	0.0089	0.0243
	(0.2111)	(0.2159)	(0.1489)	(0.1446)
Discount Frequency	0.0746	0.0879	-0.0795	-0.0787
	(0.1002)	(0.1013)	(0.0797)	(0.0797)
Herfindahl Index	0.1922	0.1971	0.1907	0.2056*
	(0.1047)	(0.1048)	(0.1064)	(0.1042)
Category Sales	-0.1027**	-0.1021**	-0.0988**	-0.0949**
	(0.0363)	(0.0360)	(0.0351)	(0.0344)
Vendor Sales	-0.0337	-0.0339	-0.0318	-0.0315
	(0.0339)	(0.0335)	(0.0336)	(0.0331)
Private Label	0.2758**	0.2725**	0.2568**	0.2494**
	(0.0445)	(0.0447)	(0.0468)	(0.0471)
Num. Customers with 1 repeats	-0.0308	-0.0262	-0.0220*	-0.0214
	(0.0238)	(0.0257)	(0.0119)	(0.0128)
Num. Customers with 2 repeats	-0.0919	-0.1000	-0.0667*	-0.0552
	(0.0591)	(0.0671)	(0.0310)	(0.0366)
Num. Customers with 3 or more repeats	0.0204	0.0332	0.0240	0.0549
	(0.0843)	(0.0913)	(0.0547)	(0.0545)
Log Likelihood	-1,851	-1,844	-1,836	-1,816
Likelihood Ratio Test, Chi^2 (df=4)		13.07*		39.15**
Area under ROC curve	0.6947	0.6985	0.7052	0.7175

The table reports average marginal effects from models where the dependent variable is a binary variable indicating whether the new product succeeded (1 if succeeded, 0 if failed). Robust standard errors (clustered at the category level) are reported in parentheses. The unit of analysis is a new product. The sample size is 2,953 new products. Significantly different from zero (or significant difference between Models 1 and 2): *p<0.05, ** p<0.01.

Table WA7: Including Products that Failed in the First 15 weeks

	Model 1	Model 2
Total Sales	0.0011**	
	(0.0004)	
Group 1 Sales		0.0118*
		(0.0048)
Group 2 Sales		0.0007
		(0.0053)
Group 3 Sales		-0.0064
		(0.0035)
Group 4 Sales		-0.0262**
		(0.0052)
Sales from Non-grouped Customers		0.0115**
		(0.0022)
Log Likelihood	-2,079	-2,031
Likelihood Ratio Test, Chi^2 (df=4)		95.19**
Area under ROC curve	0.6084	0.6158

The table reports average marginal effects from models where the dependent variable is a binary variable indicating whether the new product succeeded (1 if succeeded, 0 if failed). Robust standard errors (clustered at the category level) are reported in parentheses. The unit of analysis is a new product. The sample size is 3,106, and includes products that failed in the first 15 weeks. Significantly different from zero (or significant difference between Models 1 and 2): * $p<0.05$, ** $p<0.01$.

Table WA8: Changing the Periods Used to Divide the Classification and Prediction Sets

	First 26 Weeks for Classification		First 52 Weeks for Classification	
	Model 1	**Model 2**	**Model 1**	**Model 2**
Total Sales	0.0008** (0.0003)		0.0013** (0.0004)	
Group 1 Sales		0.0057 (0.0052)		0.0245** (0.0065)
Group 2 Sales		-0.0053 (0.0047)		0.0024 (0.0053)
Group 3 Sales		-0.0096* (0.0037)		-0.0035 (0.0040)
Group 4 Sales		-0.0166** (0.0046)		-0.0280** (0.0064)
Sales to Other Customers		0.0099** (0.0017)		0.0102* (0.0031)
Log Likelihood	-2,509	-2,461	-1,335	-1,300
Likelihood Ratio Test, Chi^2 (df=4)		96.25**		71.46**
AUC	0.5879	0.6069	0.5990	0.6300
Sample size	3,702	3,702	1,962	1,962
Observed Success Rate	42.19%	42.19%	45.16%	45.16%

The table reports average marginal effects from models where the dependent variable is a binary variable indicating whether the new product succeeded (1 if succeeded, 0 if failed). Robust standard errors (clustered at the category level) are reported in parentheses. The unit of analysis is a new product. Significantly different from zero (or significant difference between Models 1 and 2): * $p<0.05$, ** $p<0.01$.

Table WA9: Random Assignment of Products to Classification and Prediction Sets

	Randomly Assigning Products		Randomly Assigning Product Categories	
	Model 1	**Model 2**	**Model 1**	**Model 2**
Total Sales	0.0012**		0.0003	
	(0.0003)		(0.0003)	
Group 1 Sales		0.0052**		0.0249**
		(0.0052)		(0.0047)
Group 2 Sales		0.0055		-0.0059
		(0.0056)		(0.0050)
Group 3 Sales		-0.0013		-0.0105*
		(0.0032)		(0.0048)
Group 4 Sales		-0.0306**		-0.0186*
		(0.0071)		(0.0077)
Sales to Other Customers		0.0074**		0.0031
		(0.0027)		(0.0031)
Log Likelihood	-1,759	-1,685	-1,639	-1,598
Likelihood Ratio Test, Chi^2 (df=4)		148.39**		83.04**
AUC	0.5743	0.6422	0.5621	0.6143
Sample size	2,626	2,626	2,435	2,435
Observed Success Rate	40.63%	40.63%	40.12%	40.12%

The table reports average marginal effects from models where the dependent variable is a binary variable indicating whether the new product succeeded (1 if succeeded, 0 if failed). Robust standard errors (clustered at the category level) are reported in parentheses. The unit of analysis is a new product. For this analysis we restrict attention to new product purchases that are within 52 weeks of product introduction, and for which we observe the first 15 weeks of sales in our transaction data. We then randomly assign the products or the product categories into classification and prediction sets. For the prediction set, we further restrict attention to products that survive for 15 weeks. Significantly different from zero (or significant difference between Models 1 and 2): * $p<0.05$, ** $p<0.01$.

Table WA10: Cross-Market Analysis

	111 Stores Only		Not 111 Stores		Different Sample	
	Model 1	**Model 2**	**Model 1**	**Model 2**	**Model 1**	**Model 2**
Total Sales	0.0020**		0.0017**		0.0001**	
	(0.0008)		(0.0006)		(0.0000)	
Group 1 Sales		-0.0050		0.0235**		0.0004*
		(0.0082)		(0.0069)		(0.0002)
Group 2 Sales		0.0084		-0.0024		0.0008*
		(0.0080)		(0.0065)		(0.0004)
Group 3 Sales		-0.0186**		-0.0086		-0.0023*
		(0.0057)		(0.0044)		(0.0010)
Group 4 Sales		-0.0475**		-0.0183**		-0.0031
		(0.0079)		(0.0067)		(0.0028)
Sales to Other Customers		0.0189**		0.0109**		0.0013*
		(0.0028)		(0.0028)		(0.0006)
Log Likelihood	-1,841	-1,799	-1,935	-1,911	-1,581	-1,400
Likelihood Ratio Test, Chi^2 (df=4)		83.93**		48.38**		360.93**
AUC	0.5916	0.6054	0.5865	0.5872	0.7409	0.8109
Sample size	2,697	2,697	2,845	2,845	2,693	2,693
Observed Success Rate	44.38%	44.38%	43.55%	43.55%	39.21%	39.21%

The table reports average marginal effects from models where the dependent variable is a binary variable indicating whether the new product succeeded (1 if succeeded, 0 if failed). Robust standard errors (clustered at the category level) are reported in parentheses. The unit of analysis is a new product. Significantly different from zero (or significant difference between Models 1 and 2): * $p<0.05$, ** $p<0.01$.

Table WA11: Minimum Number of Classification Set Purchases

	Model 1	Model 2		
		At Least 3 Purchases	At Least 4 Purchases	At Least 5 Purchases
Total Sales	0.0011**			
	(0.0004)			
Group 1 Sales		0.0133*	0.0125**	0.0057
		(0.0063)	(0.0067)	(0.0073)
Group 2 Sales		0.0015	-0.0010	-0.0003
		(0.0054)	(0.0056)	(0.0056)
Group 3 Sales		-0.0097*	-0.0101*	-0.0107*
		(0.0041)	(0.0047)	(0.0048)
Group 4 Sales		-0.0237**	-0.0251**	-0.0293**
		(0.0059)	(0.0064)	(0.0065)
Sales to Other Customers		0.0086**	0.0079**	0.0076**
		(0.0018)	(0.0016)	(0.0015)
Log Likelihood	-1,998	-1,959	-1,959	-1,957
Likelihood Ratio Test, Chi^2 (df=4)		78.81**	78.27**	82.77**
AUC	0.6035	0.6183	0.6215	0.6235

The table reports average marginal effects from models where the dependent variable is a binary variable indicating whether the new product succeeded (1 if succeeded, 0 if failed). Robust standard errors (clustered at the category level) are reported in parentheses. The unit of analysis is a new product. The unit of analysis is a new product. The sample size is 2,953. Significantly different from zero (or significant difference between Models 1 and 2): * $p<0.05$, ** $p<0.01$.

Table WA12: Using Ratio Measures to Predict Success

	Model 1	**Model 2**
Total Sales	0.0011**	0.0011**
	(0.0004)	(0.0004)
Group 2 Sales Ratio		-0.1422*
		(0.0597)
Group 3 Sales Ratio		-0.1734**
		(0.0539)
Group 4 Sales Ratio		-0.2998**
		(0.0683)
No Sales to Grouped Customers		-0.1505**
		(0.0575)
Log Likelihood	-1,998	-1,985
Likelihood Ratio Test, Chi^2 (df=4)		27.22**
Area under ROC curve	0.6035	0.6012

The table reports average marginal effects from models where the dependent variable is a binary variable indicating whether the new product succeeded (1 if succeeded, 0 if failed). Robust standard errors (clustered at the category level) are reported in parentheses. The unit of analysis is a new product. The sample size is 2,953. Significantly different from zero (or significant difference between Models 1 and 2): *p<0.05, ** p<0.01.

Table WA13: Grouping Customers by *Success Avoidance*

	Model 1	**Model 2**
Total Sales	0.0011**	
	(0.0004)	
Group 1 Sales		0.0028
		(0.0035)
Group 2 Sales		-0.0011
		(0.0047)
Group 3 Sales		0.0052
		(0.0064)
Group 4 Sales		-0.0247**
		(0.0038)
Sales to Other Customers		0.0125**
		(0.0024)
Log Likelihood	-1,998	-1,954
Likelihood Ratio Test, Chi^2 (df=4)		87.76**
Area under ROC curve	0.6035	0.6109

The table reports average marginal effects. The dependent variable is a binary variable indicating whether the new product succeeded (1 if succeeded, 0 if failed). Robust standard errors (clustered at the category level) are reported in parentheses. The unit of analysis is a new product. The sample size is 2,953. Significantly different from zero (or significant difference between Models 1 and 2): * $p<0.05$, ** $p<0.01$.

Table WA14: Holdout Accuracy

	Model 1	Model 2
Total Sales	0.0007 (0.0004)	
Group 1 Sales		0.0096 (0.0054)
Group 2 Sales		0.0067 (0.0065)
Group 3 Sales		-0.0101* (0.0047)
Group 4 Sales		-0.0220** (0.0062)
Sales from Other Customers		0.0101** (0.0029)
Log Likelihood	-1,169	-1,148
Area under ROC curve	0.5918	0.6051
Likelihood Ratio Test, Chi^2 (df=4)		42.16**
% Correct Predictions (Holdout)	54.99%	61.42%**

The table reports average marginal effects from models where the dependent variable is a binary variable indicating whether the new product succeeded (1 if succeeded, 0 if failed). Robust standard errors (clustered at the category level) are reported in parentheses. The unit of analysis is a new product and the sample sizes for the estimation and the holdout sample are 1,740 and 1,213. The success rates for the two samples are 40.29% and 45.92%. Significantly different from zero (or significant difference between Models 1 and 2): *p<0.05, ** p<0.01.

Table WA15: Results by Super-Category

	Beauty		Edibles		General Merchandise		Healthcare	
	Model 1	**Model 2**	**Model 1**	**Model 2**	**Model 1**	**Model 2**	**Model 1**	**Model 2**
Total Sales	0.0042*		-0.0012		0.0010*		0.0020*	
	(0.0018)		(0.0006)		(0.0005%)		(0.0009)	
Group 1 Sales		0.0092		-0.0035		0.0056		0.0123
		(0.0118)		(0.0095)		(0.0074)		(0.0080)
Group 2 Sales		0.0039		-0.0048		-0.0094		0.0084
		(0.0140)		(0.0175)		(0.0076)		(0.0099)
Group 3 Sales		-0.0007		-0.0153		0.0005		-0.0099
		(0.0071)		(0.0089)		(0.0066)		(0.0070)
Group 4 Sales		-0.0200*		-0.0133		-0.0059		-0.0315**
		(0.0097)		(0.0080)		(0.0093)		(0.0102)
Sales from Non-grouped Customers		0.0139**		0.0131*		0.0051		0.0133**
		(0.0049)		(0.0058)		(0.0038)		(0.0041)
Log Likelihood	-905.5	-899.6	-103.32	-90.80	-410.72	-408.54	-517.50	-498.99
Likelihood Ratio Test, Chi^2 (df=4)		11.68*		25.04**		4.36		37.02**
AUC	0.6185	0.6258	0.7044	0.7613	0.5634	0.5643	0.6203	0.6640
Sample size	1,437	1,437	156	156	600	600	760	760

The table reports average marginal effects from models where the dependent variable is a binary variable indicating whether the new product succeeded (1 if succeeded, 0 if failed). Robust standard errors (clustered at the category level) are reported in parentheses. The unit of analysis is a new product. Significantly different from zero (or significant difference between Models 1 and 2): * $p<0.05$, ** $p<0.01$.

Table WA16: National Brand and Private Label Results

	National Brand		Private Label	
	Model 1	**Model 2**	**Model 1**	**Model 2**
Total Sales	0.0010**		-0.0001	
	(0.0003)		(0.0007)	
Group 1 Sales		0.0054		0.0147
		(0.0049)		(0.0096)
Group 2 Sales		0.0023		-0.0014
		(0.0060)		(0.0083)
Group 3 Sales		-0.0064		0.0022
		(0.0038)		(0.0057)
Group 4 Sales		-0.0236**		-0.0189
		(0.0055)		(0.0100)
Sales from Non-grouped Customers		0.0119**		0.0010
		(0.0023)		(0.0044)
Log Likelihood	-1,585	-1,553	-321.6	-317.5
Likelihood Ratio Test, Chi^2 (df=4)		63.31**		8.30
AUC	0.5856	0.6106	0.4377	0.5643
Sample size	2,431	2,431	522	522

The table reports average marginal effects from models where the dependent variable is a binary variable indicating whether the new product succeeded (1 if succeeded, 0 if failed). Robust standard errors (clustered at the category level) are reported in parentheses. The unit of analysis is a new product. Significantly different from zero (or significant difference between Models 1 and 2): * $p<0.05$, ** $p<0.01$.

Table WA17: Low vs. High-Priced Items

	Low Priced		High Priced	
	Model 1	**Model 2**	**Model 1**	**Model 2**
Total Sales	0.0008**		0.0040**	
	(0.0004)		(0.0013)	
Group 1 Sales		0.0056		0.0241**
		(0.0059)		(0.0080)
Group 2 Sales		-0.0058		0.0208
		(0.0060)		(0.0086)
Group 3 Sales		-0.0075		-0.0054
		(0.0043)		(0.0059)
Group 4 Sales		-0.0194**		-0.0286**
		(0.0060)		(0.0104)
Sales from Non-grouped Customers		0.0133**		0.0075
		(0.0026)		(0.0042)
Log Likelihood	-978.6	-946.2	-1000.9	-986.3
Likelihood Ratio Test, Chi^2 (df=4)		64.93**		29.22**
AUC	0.611	0.623	0.611	0.638
Sample size	1,467	1,467	1,486	1,486

The table reports average marginal effects from models where the dependent variable is a binary variable indicating whether the new product succeeded (1 if succeeded, 0 if failed). Robust standard errors (clustered at the category level) are reported in parentheses. The unit of analysis is a new product. Significantly different from zero (or significant difference between Models 1 and 2): * $p<0.05$, ** $p<0.01$.

Table WA18: Discount Frequency

	Less-Frequently Discounted		More-Frequently Discounted	
	Model 1	**Model 2**	**Model 1**	**Model 2**
Total Sales	0.0027**		0.0009**	
	(0.0010)		(0.0003)	
Group 1 Sales		0.0252**		0.0045
		(0.0082)		(0.0048)
Group 2 Sales		-0.0123		0.0084
		(0.0101)		(0.0055)
Group 3 Sales		-0.0005		-0.0079*
		(0.0062)		(0.0038)
Group 4 Sales		-0.0310**		-0.0190**
		(0.0089)		(0.0055)
Sales from Non-grouped Customers		0.0138**		0.0091**
		(0.0037)		(0.0024)
Log Likelihood	-997.3	-969.6	-980.7	-962.6
Likelihood Ratio Test, Chi^2 (df=4)		54.41**		36.31**
AUC	0.607	0.638	0.613	0.612
Sample size	1,465	1,465	1,488	1,488

The table reports average marginal effects from models where the dependent variable is a binary variable indicating whether the new product succeeded (1 if succeeded, 0 if failed). Robust standard errors (clustered at the category level) are reported in parentheses. The unit of analysis is a new product. Significantly different from zero (or significant difference between Models 1 and 2): * $p<0.05$, ** $p<0.01$.

Table WA19: Harbingers vs. Others Purchasing Behavior for New Products

	Harbingers	Others	Difference
Weeks After New Product Introduced	26.797	27.901	-1.104** (0.110)
Total Purchases	5.046	5.010	0.037 (0.067)
Purchases per Visit	1.384	1.280	0.105** (0.006)
Purchases per Store	3.591	3.638	-0.047 (0.047)
Shopping Visits	3.780	4.064	-0.284** (0.045)
Different Stores Visited	1.529	1.542	-0.013 (0.010)
Regular Price of Items	$5.479	$7.228	-$1.748** (0.055)
Price Paid	$4.696	$6.546	-$1.849** (0.052)
Discount Received	14.091%	10.554%	3.537%** (0.186%)
% Discounted Items	32.874%	29.044%	3.830%** (0.368%)
% Beauty Items	31.861%	23.092%	8.769%** (0.355%)
% Edible Items	10.665%	8.107%	2.559%** (0.239%)
% General Merchandise Items	25.110%	23.979%	1.130%** (0.337%)
% Health Items	32.364%	44.822%	-12.458%** (0.387%)

The table reports the purchasing behaviors for both Harbingers and Other customers. All measures are calculated using purchases of both new and existing products in the classification period of the transaction data (November 2003 to July 2004). Standard errors of the mean difference are reported in parentheses. The sample size is 29,463. Harbingers are customers from Groups 3 & 4 (n = 16,620), while others are customers from Groups 1 & 2 (n = 12,843). Significantly different from zero: * $p<0.05$, ** $p<0.01$.

Table WA20: Variable Definitions: Purchasing Behavior

Weeks after New Product Introduction	Average number of weeks the new product is bought after the introduction.
Total Purchases	Total number of purchases.
Purchases per Visit	Average number of purchases per shopping visit.
Purchases per Stores	Average number of purchases per store.
Shopping Visits	Total number of shopping visits.
Different Stores Visited	Total number of different stores visited.
Regular Price of Items	Average regular retail price of items bought.
Price Paid	Average price paid of items bought.
Discount Received	Average promotion depth of discounted items bought, calculated as the average of: (regular price – price paid)/regular price.
% Discounted Items	Percentage of items bought at discount.
% Beauty Items	Percentage of beauty items, such as skin care, hair care, etc.
% Edible Items	Percentage of edible items, such as beverages, snacks, etc.
% General Merchandise Items	Percentage of general merchandise items, such as stationery, housewares, etc.
% Health Items	Percentage of health items, such as eye care, cold remedies, etc.

These measures are all calculated using purchases in the classification period of the transaction data (November 2003 – July 2004).

Table WA21: Summary Statistics: Purchasing Behavior

	All Products		New Products	
	Mean	**Standard Deviation**	**Mean**	**Standard Deviation**
Weeks after New Product Introduction			27.278	9.339
Total Purchases	74.768	87.038	5.031	5.732
Purchases per Visit	3.906	2.253	1.339	0.554
Purchases per Store	34.478	41.232	3.612	4.004
Shopping Visits	20.539	21.002	3.904	3.868
Different Stores Visited	2.828	2.564	1.535	0.888
Regular Price of Items	$4.559	$1.848	$6.242	$4.720
Price Paid	$3.919	$1.771	$5.503	$4.484
Discount Received	13.793%	10.509%	12.549%	15.887%
% Discounted Items	34.405%	19.777%	31.205%	31.358%
% Beauty Items	17.485%	14.503%	28.039%	30.523%
% Edible Items	29.815%	17.423%	9.550%	20.385%
% General Merchandise Items	21.957%	13.337%	24.617%	28.710%
% Health Items	23.559%	15.680%	37.794%	33.500%

The table reports the average and standard deviations of the demographic measures for the 29,463 classified customers.