

**Online Appendix A. Proof that the Profit-Maximizing Policy is a Threshold Policy, that the Threshold Policy is Intuitive, and that Profits are Decreasing in Predictive Uncertainty.**

**Result 1.** Suppose (1) the firm's prior on the profitability of an item,  $\pi$ , is normally distributed,  $\pi \sim \mathcal{N}(\mu_0; \sigma_0^2)$ , (2) the firm observes an estimate of profitability  $\hat{\pi}|\pi \sim \mathcal{N}(\pi; \sigma_1^2)$ , and (3) the firm seeks a policy to decide whether to put an item online or not. Then the profit maximizing policy,  $\mathcal{P}(\phi)$ , is a threshold policy:

$$(A1) \quad \mathcal{P}(\phi) = \begin{cases} 1 & \text{if } \hat{\pi} \geq -\mu_0 \frac{\sigma_1^2}{\sigma_0^2} \\ 0 & \text{if } \hat{\pi} < -\mu_0 \frac{\sigma_1^2}{\sigma_0^2} \end{cases}$$

The policy in Equation A1 is intuitive. For example,

- If predictions are perfect, then  $\sigma_1^2 = 0$  and the policy reverts to that of perfect prediction; launch those items for which  $\hat{\pi} \geq 0$ .
- If the model has no predictive ability, then  $\sigma_1^2 \rightarrow \infty$  and the policy reverts to the prior mean,  $\mu_0$ ; launch all items if and only if the prior mean is positive.
- If there is no uncertainty in the prior, then  $\sigma_0^2 \rightarrow 0$  and the policy again reverts to the prior mean; launch all items if and only if the prior mean is positive.
- For finite values of  $\sigma_1^2$  and  $\sigma_0^2$ , the ratio,  $\sigma_1^2/\sigma_0^2$ , modifies the amount by which the predicted profits must exceed prior beliefs in order to launch.

**Proof of the threshold policy:** The firm solves the following optimization problem:

$$(A2) \quad \max_{\mathcal{P}(\phi) \in [0,1]} \mathbb{E}[\mathcal{P}(\phi) * \pi + (1 - \mathcal{P}(\phi)) * 0] = \max_{\mathcal{P}(\phi) \in [0,1]} \mathbb{E}[\mathcal{P}(\phi) * \pi]$$

where  $\phi \equiv (\hat{\pi}, \mu_0, \sigma_0^2, \sigma_1^2)$  is the set of all known parameters;  $\hat{\pi}|\pi \sim \mathcal{N}(\pi; \sigma_1^2)$  and  $\pi \sim \mathcal{N}(\mu_0; \sigma_0^2)$

Using the law of iterative expectations, we rewrite the initial maximization problem (A2) as:

$$(A3) \quad \max_{\mathcal{P}(\phi) \in [0,1]} \mathbb{E}[\mathcal{P}(\phi) * \pi] = \max_{\mathcal{P}(\phi) \in [0,1]} \mathbb{E}[\mathcal{P}(\phi) * \mathbb{E}[\pi|\phi]] = \max_{\mathcal{P}(\phi) \in [0,1]} \mathbb{E}[\mathcal{P}(\phi) * \mathbb{E}[\pi|\hat{\pi}]]$$

The last step relies on the assumption that  $\sigma_0, \sigma_1, \mu_0$  are observable.

Because  $\mathbb{E}[\pi|\phi]$  is a function of observables,  $\phi$ , we can denote  $\mathbb{E}[\pi|\phi] = f(\phi)$ . Equation (A3) is rewritten as:

$$(A4) \quad \max_{\mathcal{P}(\phi) \in [0,1]} \mathbb{E}[\mathcal{P}(\phi) * f(\phi)]$$

Equation (A4) implies that the optimal policy  $\mathcal{P}^*(\phi)$  has the following form ( $\mathcal{J}(\cdot)$  is an indicator function):

$$(A5) \quad \mathcal{P}^*(\phi) = \mathcal{J}(f(\phi) \geq 0) = \mathcal{J}(\mathbb{E}[\pi|\phi] \geq 0)$$

We show in the following that, for the case of normal priors, this policy would have a threshold form. [Note that the optimal policy in Equation (A5) does not depend on the normality assumption profitability; the policy is easily generalized to other distributions.]

Because  $\hat{\pi}$  is normally distributed conditionally on  $\pi$  and since the prior is also normally distributed, the posterior is normally distributed. Using standard formulae, we write:

$$(A6) \quad \pi|\hat{\pi} \sim \mathcal{N}\left(\frac{\hat{\pi}\sigma_0^2 + \mu_0\sigma_1^2}{\sigma_0^2 + \sigma_1^2}; \frac{\sigma_0^2\sigma_1^2}{\sigma_0^2 + \sigma_1^2}\right) \quad \text{and} \quad \hat{\pi} \sim \mathcal{N}(\mu_0; \sigma_0^2 + \sigma_1^2)$$

From (A6), it follows that:

$$(A7) \quad \mathbb{E}[\pi|\phi] = \frac{\hat{\pi}\sigma_0^2 + \mu_0\sigma_1^2}{\sigma_0^2 + \sigma_1^2} \Rightarrow \mathcal{P}^*(\phi) = \mathcal{J}(\mathbb{E}[\pi|\phi] \geq 0) = \mathcal{J}\left(\hat{\pi} \geq -\mu_0 * \frac{\sigma_1^2}{\sigma_0^2}\right)$$

Which is the threshold policy.

**Result 2.** Under the assumptions of Result 1, the optimal expected profit is:

$$(A8) \quad \Pi^* = \left(1 - \Phi\left(-\frac{\mu_0}{\sigma_v}\right)\right) * \mu_0 + \sigma_v * \varphi\left(-\frac{\mu_0}{\sigma_v}\right)$$

Where  $\Phi(\cdot)$  and  $\varphi(\cdot)$  are the standard normal CDF and PDF respectively, and  $\sigma_v = \frac{\sigma_0^2}{\sqrt{\sigma_0^2 + \sigma_1^2}}$ .

**Proof:** By substituting the optimal policy from (A7) and conditional expectation from (A8) to (A2), we rewrite the expected optimal profit as:

$$(A9) \quad \Pi^* = \mathbb{E}\left[\mathcal{J}\left(\hat{\pi} \geq -\mu_0 * \frac{\sigma_1^2}{\sigma_0^2}\right) * \left(\frac{\hat{\pi}\sigma_0^2 + \mu_0\sigma_1^2}{\sigma_0^2 + \sigma_1^2}\right)\right] = \mathbb{E}[\mathcal{J}(v \geq 0) * v] = \mathbb{P}[v \geq 0]\mathbb{E}[v|v \geq 0]$$

where  $v = \frac{\hat{\pi}\sigma_0^2 + \mu_0\sigma_1^2}{\sigma_0^2 + \sigma_1^2} \sim \mathcal{N}\left(\frac{\hat{\pi}\sigma_0^2 + \mu_0\sigma_1^2}{\sigma_0^2 + \sigma_1^2}; \frac{\sigma_0^4}{(\sigma_0^2 + \sigma_1^2)^2}(\sigma_0^2 + \sigma_1^2)\right) \sim \mathcal{N}\left(\mu_0; \frac{\sigma_0^4}{\sigma_0^2 + \sigma_1^2}\right) \sim \mathcal{N}(\mu_0; \sigma_v^2)$

Because  $v$  is normally distributed, (A9) can be rewritten using the formula for the expectation of the truncated normal distribution:

$$(A10) \quad \Pi^* = \left(1 - \Phi\left(-\frac{\mu_0}{\sigma_v}\right)\right) * \mu_0 + \sigma_v * \varphi\left(-\frac{\mu_0}{\sigma_v}\right)$$

**Result 3.** The expected profit under the optimal policy is a decreasing function of  $\sigma_1^2$ .

**Proof:** Taking the derivative of (A10) with respect to  $\sigma_1^2$ :

$$(A11) \quad -\mu_0 * \varphi\left(-\frac{\mu_0}{\sigma_v}\right) \left(-\frac{\mu_0}{2\sigma_0^2(\sigma_0^2 + \sigma_1^2)^{\frac{1}{2}}}\right) - \frac{\sigma_0^2}{2(\sigma_0^2 + \sigma_1^2)^{\frac{3}{2}}} \varphi\left(-\frac{\mu_0}{\sigma_v}\right) +$$

$$\frac{\sigma_0^2}{(\sigma_0^2 + \sigma_1^2)^{\frac{1}{2}}} \varphi'\left(-\frac{\mu_0}{\sigma_v}\right) \left(-\frac{\mu_0}{2\sigma_0^2(\sigma_0^2 + \sigma_1^2)^{\frac{1}{2}}}\right) = \left(\frac{\mu_0^2}{2\sigma_0^2(\sigma_0^2 + \sigma_1^2)^{\frac{1}{2}}} - \frac{\sigma_0^2}{2(\sigma_0^2 + \sigma_1^2)^{\frac{3}{2}}}\right) +$$

$$\frac{\sigma_0^2}{(\sigma_0^2 + \sigma_1^2)^{\frac{1}{2}}} \left(\frac{\mu_0(\sigma_0^2 + \sigma_1^2)^{\frac{1}{2}}}{\sigma_0^2}\right) \left(-\frac{\mu_0}{2\sigma_0^2(\sigma_0^2 + \sigma_1^2)^{\frac{1}{2}}}\right) \varphi\left(-\frac{\mu_0}{\sigma_v}\right) = \left(\frac{\mu_0^2}{2\sigma_0^2(\sigma_0^2 + \sigma_1^2)^{\frac{1}{2}}} - \frac{\sigma_0^2}{2(\sigma_0^2 + \sigma_1^2)^{\frac{3}{2}}}\right) +$$

$$\left(-\frac{\mu_0^2}{2\sigma_0^2(\sigma_0^2 + \sigma_1^2)^{\frac{1}{2}}}\right) \varphi\left(-\frac{\mu_0}{\sigma_v}\right) = -\frac{\sigma_0^2}{2(\sigma_0^2 + \sigma_1^2)^{\frac{3}{2}}} \varphi\left(-\frac{\mu_0}{\sigma_v}\right)$$

Because  $\varphi(\cdot) > 0$  and  $-\frac{\sigma_0^2}{2(\sigma_0^2 + \sigma_1^2)^{\frac{3}{2}}} < 0$ , the expected profitability is decreasing function of  $\sigma_1^2$  and

therefore an increasing function of model accuracy.

## Online Appendix B. Supporting Tables and Figures

**Table B.1.** Improvement in Predictive Accuracy Varying Minimum Threshold on Online Sales

Model	Non-Image Features	Image Features	Out of Sample R <sup>2</sup>	Improvement over baseline
CNN Features with <b>10</b> as threshold for online sales	Category, seasonality, price, color labels	Deep-learning	43.14 (0.20)	+12.75%
CNN Features with <b>20</b> as threshold for online sales	Category, seasonality, price, color labels	Deep-learning	46.88 (0.19)	+13.48%
CNN Features with <b>30</b> as threshold for online sales	Category, seasonality, price, color labels	Deep-learning	51.23 (0.17)	+12.08%

Note: Improvements are calculated for baseline models estimated on the corresponding samples. Standard deviations are reported in parentheses.

**Table B.2.** Tests of Uniqueness, Precision (variance of  $N_{i,purchased}$ ), and Distance from Prior Collections

Model	Non-Image Features	Image Features	Out of Sample R <sup>2</sup>	Improvement over CNN Features Model
CNN Features	Category, seasonality, price, color labels	Deep-learning	46.88 (0.19)	0.00%
CNN Features (including uniqueness)	Category, seasonality, price, color labels, image uniqueness	Deep-learning	46.82 (0.23)	-0.13%
CNN Features (including vs. last year)	Category, seasonality, price, color labels, image uniqueness	Deep-learning	44.55 (0.39)	-0.60% (see note)
CNN Features (precision weighting)	Category, seasonality, price, color labels, variance weighting	Deep-learning	46.77 (0.26)	-0.23%

Notes: The sample of items included when estimating the last-year model exclude products sold only in the first year of the data. A GBRT/CNN model for the same items yields 44.85 (0.33). The -2.8% is relative to this model. Standard deviations are reported in parentheses.

**Table B.3.** Improvement in Predictive Accuracy Using Alternative Prediction Models

Model	Non-Image Features	Image Features	Out of Sample R <sup>2</sup>	Improvement over Baseline
GBRT (CNN Features)	Category, seasonality, price, color labels	Deep-learning	46.88 (0.19)	+13.48%
Bagging Methods (CNN Features)	Category, seasonality, price, color labels	Deep-learning	45.35 (0.14)	+9.78%
LASSO (CNN Features)	Category, seasonality, price, color labels	Deep-learning	44.11 (0.32)	+6.78%

Note: Standard deviations are reported in parentheses.

**Table B.4.** Predictions for the two Largest Categories (Dresses and Shirts)

Model	Non-Image Features	Image Features	Out of Sample R <sup>2</sup>	Improvement over Baseline
Non-Image Baseline	Category, seasonality, price	None	57.94 (0.27)	–
Color-labels	Category, seasonality, price, and color labels	None	59.79 (0.24)	+3.19%
Automated Color Features	Category, seasonality, price, color labels	RGB	60.86 (0.22)	+5.04%
Automated Color and Patterns	Category, seasonality, price, color labels	RGB + Gabor	61.91 (0.31)	+6.85%
Human-coded features	Category, seasonality, price, color labels, human-coded features	Human-coded	61.91 (0.27)	+6.85%
CNN Features	Category, seasonality, price, color labels	Deep-learning	63.69 (0.19)	+9.92%

Note: Standard deviations are reported in parentheses.

**Table B.5.** Improvement in Predictive Accuracy Using an Alternative CNN

Model	Non-Image Features	Image Features	Out of Sample R <sup>2</sup>	Improvement over Baseline
ResNet CNN (this paper)	Category, seasonality, price, color labels	Deep-learning	46.88 (0.19)	+13.48%
VGG-19 CNN	Category, seasonality, price, color labels	Deep-learning	46.84 (0.18)	+13.39%

Note: Standard deviations are reported in parentheses.

**Table B.6.** Improvement in Predictive Accuracy Using PCA (nonlinear and linear tested; linear shown)

Model	Non-Image Features	Image Features	Out of Sample R <sup>2</sup>	Improvement over Baseline
Color Features	Category, seasonality, price, color labels	RGB	43.48 (0.18)	+5.25%
Color and Patterns	Category, seasonality, price, color labels	Gabor	41.37 (0.33)	+0.15%
CNN Features	Category, seasonality, price, color labels	Deep-learning	46.55 (0.21)	+12.68%

Note: Standard deviations are reported in parentheses.

**Table B.7.** Predictions Using Automated Pattern & Color Image-Processing Features

Model	Non-Image Features	Image Features	Out of Sample R <sup>2</sup>	Improvement over Baseline
Non-Image Baseline	Category, seasonality, price	None	41.31 (0.18)	–
Color Features	Category, seasonality, price, color labels	RGB	44.06 (0.20)	+6.66%
Pattern Features	Category, seasonality, price, color labels	Gabor	44.34 (0.23)	+7.33%
Color and Patterns	Category, seasonality, price, color labels	RGB + Gabor	45.28 (0.18)	+9.61%
CNN Features	Category, seasonality, price, color labels	Deep-learning	46.88 (0.19)	+13.48%
CNN Features all images	Category, seasonality, price, color labels	Deep-learning	47.48 (0.22)	+14.93%

Notes: All models use LightGBM and differ only with respect to the set of features included. Standard deviations are reported in parentheses.

**Table B.8.** Improvement in Predictive Accuracy Using Alternative Image-Feature Extraction Methods

Model	Non-Image Features	Image Features	Out of Sample R <sup>2</sup>	Improvement over Baseline
RGB Features	Category, seasonality, price, color labels	RGB	44.04 (0.20)	+6.61%
HSV Features	Category, seasonality, price, color labels	HSV	44.30 (0.14)	+7.24%
ORB Features	Category, seasonality, price, color labels	ORB	43.53 (0.25)	+5.37%

Note: Standard deviations are reported in parentheses.

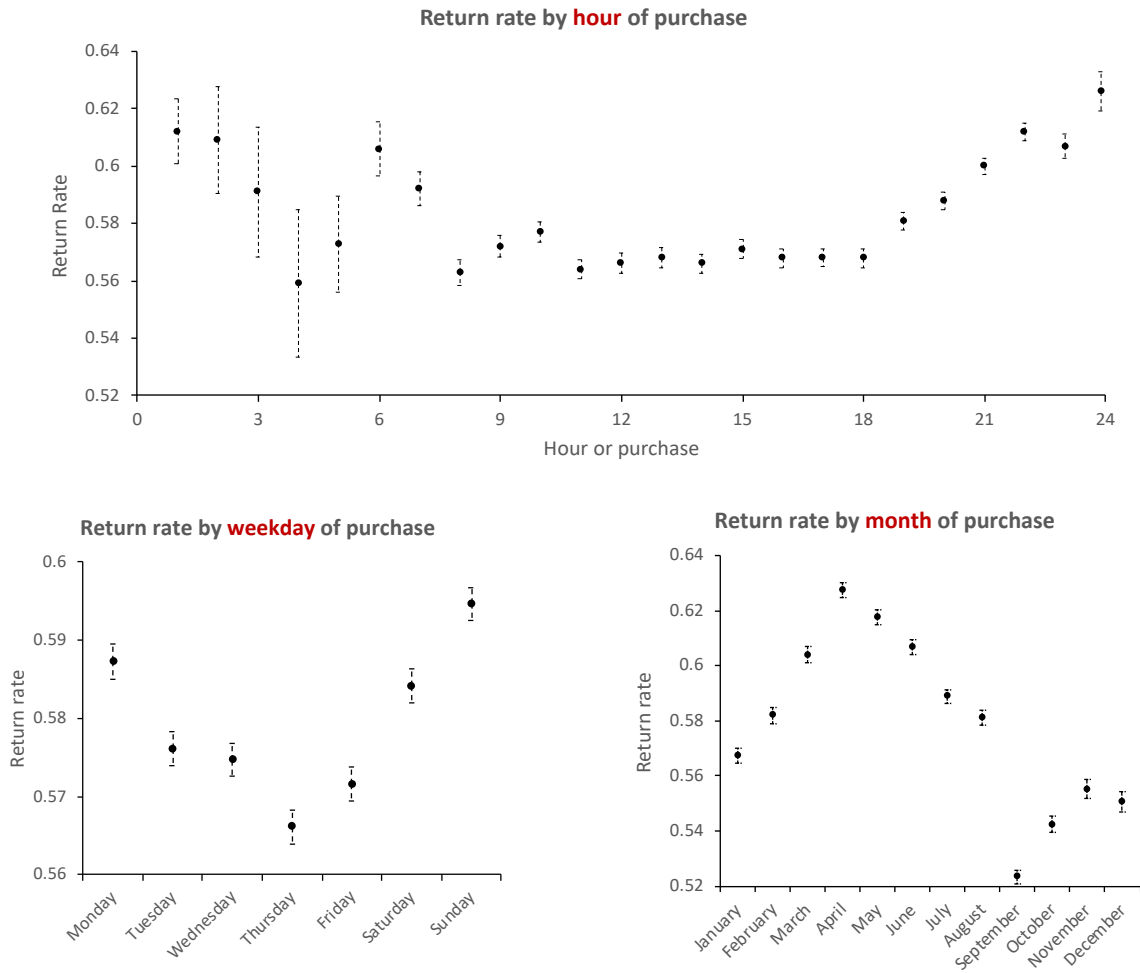
**Table B.9.** Online Sales and Return rates, Offline Sales and Return Rates, and Model-Predicted Online Return Rates by Product Category (based on all sales). Models estimated for categories with at least 400 items with  $\geq 20$  sales.

Category	Online Sales	Online Returns	Online Return Rate	Offline Sales	Offline Returns	Offline Return Rate	Number of Products $\geq 20$ Sales	Predictive Accuracy Online, Out of Sample R <sup>2</sup> (Benchmark)	Predictive Accuracy Online, Out of Sample R <sup>2</sup> (Main)
Dresses	96,754	69,626	71.96%	45,923	1,615	3.52%	759	28.11 (0.65)	31.13 (0.83)
Shirts	80,586	39,379	48.87%	299,313	7,007	2.34%	1,213	14.78 (0.59)	24.08 (0.77)
Blouses	43,413	23,292	53.65%	104,778	2,667	2.55%	687	4.33 (0.96)	15.78 (1.00)
Pants	36,183	21,209	58.62%	103,353	3,264	3.16%	496	-1.10 (1.16)	-0.17 (1.25)
Knit	31,893	15,708	49.25%	137,227	3,889	2.83%	511	1.80 (1.24)	17.35 (1.13)
Jackets	21,304	12,228	57.40%	24,385	876	3.59%	302		
Blazer	13,190	7,627	57.82%	27,748	993	3.58%	166	-	-
Cardigans	11,315	4,167	36.83%	16,462	507	3.08%	69	-	-
Skirts	9,252	5,259	56.84%	26,884	746	2.77%	135	-	-
Coats	5,170	3,238	62.63%	1,299	49	3.77%	88	-	-
Bolero	4,867	3,367	69.18%	0	0	0.00%	41	-	-
Sweatshirts	3,862	2,170	56.19%	6,126	191	3.12%	56	-	-
Jumpsuits	1,902	1,303	68.51%	462	10	2.16%	27	-	-
Top	1,543	728	47.18%	0	0	0.00%	27	-	-
Leather	614	287	46.74%	809	34	4.20%	8	-	-

**Table B.10.** Online Sales and Return rates, Offline Sales and Return Rates, and Model-Predicted Online Return Rates by Color Labels† Models estimated for color-label categories with at least 400 items with  $\geq 20$  sales.

Color Category	Online Sales	Online Returns	Online Return Rate	Offline Sales	Offline Returns	Offline Return Rate	Number of Products $\geq 20$ Sales	Predictive Accuracy Online, Out of Sample $R^2$ (Benchmark)	Predictive Accuracy Online, Out of Sample $R^2$ (Main)
Blue	84,947	49,054	57.75%	217,457	5,658	2.60%	1056	45.28 (0.28)	50.06 (0.47)
Grey	67,381	39,320	58.35%	92,119	2,787	3.02%	951	34.06 (0.58)	36.36 (0.54)
White	48,751	26,913	55.21%	118,060	3,154	2.67%	743	23.25 (0.60)	23.79 (0.87)
Red	42,708	25,749	60.29%	81,625	2,294	2.81%	542	45.80 (0.62)	50.90 (0.79)
Brown	41,540	23,557	56.71%	84,227	2,446	2.90%	590	19.09 (0.77)	24.72 (0.95)
Black	32,444	19,305	59.50%	67,663	2,028	3.00%	411	41.01 (0.53)	42.57 (0.83)
Green	10,550	6,581	62.38%	19,815	466	2.35%	144	-	-
Pink	3,539	2,118	59.85%	8,588	249	2.90%	53	-	-
Orange	3,054	1,865	61.07%	7,701	202	2.62%	53	-	-
Yellow	1,670	1,017	60.90%	2,356	64	2.72%	23	-	-
Violet	938	617	65.78%	2,273	66	2.90%	14	-	-
Several	318	151	47.48%	84,314	2224	2.66%	5	-	-

**Figure B.11.** Return Rate by Postlaunch Time of Purchase, Day of the Week and Month



**Table B.12.** Product Return Rates and Price Discounts

<b>Dependent Measure: Return rate</b>				
	<b>Model 1</b>	<b>Model 2</b>	<b>Model 3</b>	<b>Model 4</b>
Proportion discounted	0.080*** (0.008)	0.086*** (0.007)	0.088*** (0.007)	0.078*** (0.007)
Price (log10)	0.228*** (0.012)	0.221*** (0.012)	0.214*** (0.012)	0.279*** (0.008)
Intercept	-0.156*** (0.044)	-0.113*** (0.041)	-0.083** (0.041)	0.028** (0.015)
Category Controls	Yes	Yes	Yes	No
Color Controls	Yes	Yes	No	No
Seasonality Controls	Yes	No	No	No
# observations	4585	4585	4585	4585
Adjusted R-squared (model)	0.434	0.414	0.403	0.258

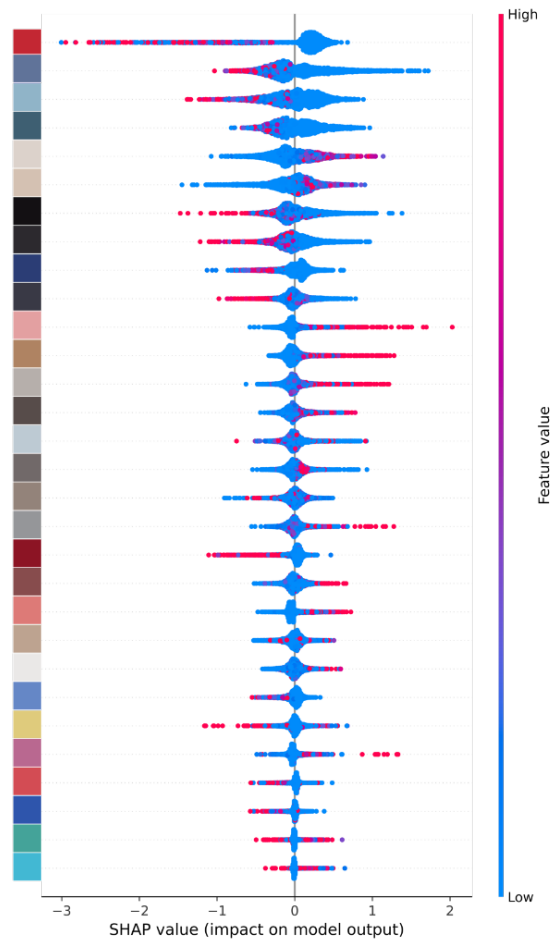
Note: Standard errors are heteroskedasticity robust (\* $p \leq 0.1$ , \*\* $p \leq 0.05$ , \*\*\* $p \leq 0.01$ ). R-squared is in-sample

**Table B.13.** Interpreting the Effect of Human-coded features (HCF) on Item Return Rates

	Regression Model	SHAP Values
Asymmetric	0.022*** (0.008)	0.92
Floral	-0.038*** (0.010)	-0.90
Striped	-0.063*** (0.009)	-0.95
Geometric/abstract	-0.020*** (0.007)	-0.89
Lace details	0.010 (0.008)	0.85
Metallic/sequin details	0.008 (0.006)	0.82
Graphic details	0.008 (0.013)	0.34
Text details	0.036* (0.021)	0.69
Short Sleeves	-0.020*** (0.006)	-0.81
Medium Sleeves	-0.032*** (0.008)	-0.84
Long Sleeves	-0.032*** (0.007)	-0.87
Belt	0.025*** (0.010)	0.88
Zipper	-0.027** (0.012)	-0.80
Intercept	0.039 (0.026)	–
Price (log10)	0.256*** (0.014)	–
Category Controls	Yes	Yes
Color Controls	Yes	Yes
Seasonality Controls	Yes	Yes
# observations	1,972	1,972
Adjusted R-squared (model)	0.628	–

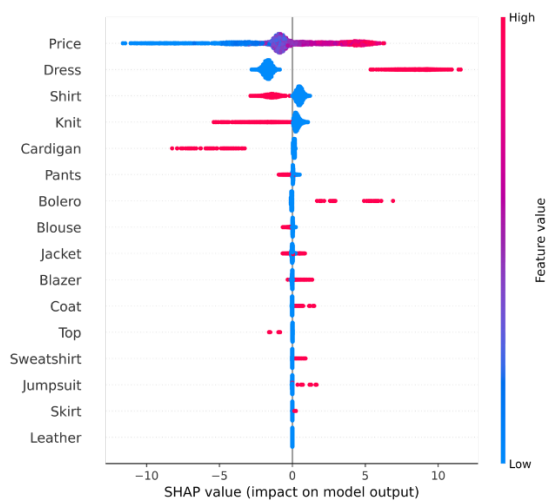
Notes: Standard errors are heteroskedasticity robust (\* $p \leq 0.1$ , \*\* $p \leq 0.05$ , \*\*\* $p \leq 0.01$ ). Products included if sales  $\geq 20$ .

**Figure B.14.** Impact of Color Clusters on Return Rates



**Figure B.15.** The Impact of Non-Image Features

(a) Price and Category



(b) Seasonality

