Can Fast Fashion Increase the Demand for Premium Brands? A Structural Analysis

Zijun (June) Shi, Param Vir Singh, Dokyun Lee, Kannan Srinivasan

(Preliminary draft. Please do not cite without the authors’ permission.)

Abstract

Erosion of high-end fashion brands by fast fashion copycats (e.g., Zara, Forever 21) has raised attention from industry insiders and policymakers, demonstrated by unceasing legal attempts to try to get fashion designs covered by copyright. This paper develops a dynamic structural model of fashion choices of brands and styles to investigate the implication of prohibiting fashion copycats, with the help of user-generated data from fashion-specific social media and deep learning methods on image analytics.

We find that copycats can enhance high-end brands demand, contrary to conventional wisdom, due to several novel mechanisms: first, the affordability of mixing low-end copycats with high-end brands boosts demand for high-end brands from financially constrained consumers.; second, good styles from low-end brands can help a consumer to build up his popularity/likeability, which increases his value for high-end brands and reduces the cost.

Substantively, our results shed light on copyright enforcement and have implications on how fashion brands should react to copycats. Methodically, we developed a framework to analyze consumer choices where visual features are important product attributes and peer feedback hugely affects the decision-making process.
1. Introduction

The global fashion industry has reached an estimated value of 3 trillion dollars.\footnote{https://fashionunited.com/global-fashion-industry-statistics} Traditional luxury fashion brands, such as Gucci, Prada, and Louis Vuitton, have maintained a strong position within the industry, backstopped by the increasing demand from developing economies such as China. At the same time, fast fashion brands such as Zara, Forever 21, and H&M have been storming the globe with their versatile styles and low price. The path to success of these fast fashion brands, however, is nothing short of controversy. Every year, large fast fashion chains spew close-to-the-runway originals at lightning speed. On the one hand, the high-end brands, believing these copycats will steal their customers and hurt their profitability, spare no effort in fighting back by launching lawsuits against them.\footnote{For example, https://wwd.com/business-news/legal/the-5-five-biggest-lawsuits-facing-fashion-retail-10875211/} On the other hand, high-end brands may not face any threat if their consumers and those of fast fashion brands are different segments with variant values of brands and styles. In addition, lawmakers tend to view the utilitarian nature of clothing and fashion as more important than its artistic and stylistic purposes; therefore fashion designs are not under the protection of copyright law.\footnote{http://www.thefashionlaw.com/home/how-do-fast-fashion-retailers-get-away-copying-high-fashion-brands} In spite of this tension on the enforcement and effects of copycats, the effect of fashion copycats on high-end brands remains empirically unclear.

In this paper, we estimate the impact of low-end copycats on the demand for high-end brands. We develop a dynamic structural model of individual consumer’s fashion choices, which allows for counterfactual analysis of alternative copyright policies against copycats. Contrary to the conventional wisdom, we find that prohibiting low-end copycats can decrease the demand of high-end brands significantly. We find novel mechanisms contributing to this result, which are distinct from the promotional effect documented in the counterfeit literature (e.g., Qian, 2014): first, fewer style choices from low-end brands would limit the mix-and-match choices for consumers and put them on greater financial constraint to get a satisfactory ensemble of clothes, resulting in them buying less high-end brands; second, the lack of good styles from low-end brands will make it harder for consumers to build up their popularity/likeability, which limits the

---

1 https://fashionunited.com/global-fashion-industry-statistics
complementary value of high-end brands. As a result, consumers adopt less high-end brands. Our findings suggest that the above-mentioned market expansion effect dominates the competition effect. Other counterfactual analyses examine the consequence if the brand or peer feedback cannot be seen, which is the case for many other social medium and offline markets.

We overcome substantial technical and empirical challenges to obtain these results. Traditionally, there are two challenges to studying the micro-level consumer choices of fashion goods. First, fashion styles are not quantifiable. Second, individual-level choices on fashion brands and styles across a large pool of brands are not available. In this paper, we employ state-of-the-art deep learning techniques to quantify fashion styles from fashion images. We overcome the second challenge by studying the choices over brands and styles for fashion conscious consumers on social media. Nowadays, fashion is one of the most popular contents generated by users on social media (Hu et al. 2014). More and more fashion consumers post what they wear online, and importantly, these social media users become trendsetters and influence a large number of other fashion consumers. Therefore, investigating how these consumers make choices on brand and style can help us understand the market demand of fashion goods. Our data is from a large online fashion-sharing platform where users post their fashion pictures and evaluate others’ pictures. The data comprise 10262 active users and 64681 fashion posts and span over three years. We account for consumer heterogeneity and estimate the structural model following a Hierarchical Bayesian framework.

Substantively, our results have implementable policy implications to both managers as well as policymakers. Managerially, we provide novel insights on how copycats can help the high-end brands, which guide their product strategy. In fact, some high-end brands have started to produce their own low-end frugal version of similar styles, consistent with the first mechanism of copycat effects in our findings. Moreover, for fashion companies, understanding how fashion consumers value brands and styles can help managers infer the market demand and make the optimal investments in branding and product design. From the policy-making perspective, we provide novel insights on the potential consequence of alternative copyright policies for fashion designs. More generally, our findings speak to the debate on whether copyright or patent protection

\[\text{http://www.latimes.com/fashion/la-ig-bloggers-20160809-snap-story.html}\]
encourages or discourages innovation in the fashion industry: with more demand brought by low-end copycats, companies can get more money to invest in innovation, which may lead to more creative designs for the entire fashion market.

This paper also contributes methodologically and theoretically. Methodologically, we make two contributions. First, we develop a framework to analyze consumer choices where visual features are important product attributes and other people’s opinions heavily affect the decision-making. Second, we use deep learning and image processing techniques to quantify fashion styles to make the analysis of fashion style choices possible. Theoretically, our findings provide new insights on how copycat products can benefit the original brands, which also apply to the cases of counterfeits and pirated goods if consumers have mix-and-match choices and popularity concern.

The rest of the paper is organized as follows. In Section 2, we review the literature related to this paper. Section 3 presents the raw data, visual feature extraction, and exploratory analysis. Section 4 describes our model. We illustrate the identification and estimation strategy in Section 5. We report the estimated results in Section 6, followed by the counterfactual analysis in Section 7. Section 8 concludes.

2. Literature Review

Our study relates to marketing and economics literature on branding, counterfeits and piracy, conspicuous good consumption, as well as the literature on machine learning methods and applications.

As we seek to examine the brand value and style value for fashion goods, our paper is related to the marketing literature on branding (e.g., Borkovsky et al., 2017; Goldfard, 2009; Keller & Lehmann, 2006; Kamakura & Russell, 1993). More recently with unstructured data, Nam et al. (2017) investigate the qualitative brand information harvested from social tags in the textual form. Liu et al. (2018) study consumers’ brand perception on social media using visual data. In this paper, we focus on fashion goods, specifically clothing. We examine how consumers value brands versus styles, and how copycat styles affect the demand of premium brands measured by the units of clothing items adopted in the social media posts.
This paper relates to the literature of counterfeits and piracy (e.g., Qian, 2014; Ma et al., 2016; Oberholzer-Gee & Strumpf, 2007; Smith & Telang, 2009), which provided evidences of both cannibalization and promotional effects of counterfeits (pirated goods) on the original. In contrast, copycats and counterfeits are fundamentally different. Counterfeits copy not only the style or content but also the trademark (i.e., the brand logo), therefore they violate the trademark law. However, copycats do not copy the brand logo and are typically legal. Therefore, counterfeits can benefit the original brand by improving the awareness of the brand (i.e., promotional effect), but copycats cannot directly give consumers information about the original brands. Studies that examine the market response of copycats include Horen and Pieters (2012) and Wang et al. (2018). Horen and Pieters (2012) conducted lab experiments and survey studies at a grocery context to demonstrate how copycats can gain or lose from their resemblance to the original brands, but they remain silent about how copycats affect the demand of the original brands. Wang et al. (2018) examine the aggregate impact of copycat mobile apps on the demand of original apps. They find that deceptive and low-quality copycat apps may positively affect the demand of the original app, implying the existence of the promotional effect. In contrast, fashion goods are fundamentally different from grocery and mobile apps, in the sense that consumers mix and match multiple clothing items and peer feedback plays an important role. Moreover, as will be shown in this paper, our micro-level study specifies new mechanisms on how copycats can benefit premium brands, unlike the traditional promotional effect.

Various theoretical works have investigated fashion firms’ strategies on information disclosure (e.g., Yoganarasimhan, 2012), competitive pricing (e.g., Amaldoss & Jain, 2005), given consumers’ dual needs of conformity and differentiation in conspicuous consumption (Brewer, 1991). Accordingly, the consumer tradeoff between expressing individuality and conforming to others’ opinions/likes, related to self and public self-consciousness for social behaviors in psychology (Fenigstein et al., 1975). Though people have heterogeneous underlying preferences, Bernheim (1994) shows that when status is very important relative to intrinsic utility, some people will conform to a single standard of behavior. To our knowledge, this is the first empirical research examining consumers’ conspicuous consumption while incorporating both their intrinsic preferences and the impact of others’ opinions.

To extract and quantify the styles of fashion goods, we need to analyze visual data by referring to machine learning literature for image analytics. Specifically, we apply support vector machine
(SVM), support vector regression (SVR), Fast R-CNN (Girshick, 2015), and transfer learning using Siamese CNN (Hadsell et al., 2006; Veit et al., 2015) to extract clothing style features (i.e., compatibility and distinctiveness) and user appearance features (i.e., facial attractiveness and body BMI).

3. Data

The research context in this paper is the world’s largest online fashion sharing community, designed for users to post their own fashion photography, featuring themselves and their outfits. It shares similar features with other photo-based social media except that the content is restricted to fashion. More importantly, the website features a special function: a user tags the brand of the fashion items in his/her posted picture. Therefore, the brand information is clearly listed beside the fashion look and can be seen by others. Figure 1 shows what a fashion post looks like on the website.

Figure 1. Example of a Fashion Post

---

5 The company was launched in 2008. As of July 2017, there are more than 6 million users registered.
6 We use “fashion look” to refer to a picture.
We collect individual-level historical data from August 2013 to August 2017. The data set contains the entire history of fashion content generation for a random sample of 10,262 users who registered after August 2013 and posted at least once. For each fashion look, we collect the image data, the brands for the clothing items, the time stamp, as well as how many likes the picture has attracted. For each user, we also observe his/her age from the brief biography. The gender information is not directly observable, and we will predict the gender from their picture in section 3.1.2.

Table 1 shows the summary statistics for these users and their fashion posts. We can see that the standard deviation is large relative to the means, and the measures have skewed distribution. A small group of people has lots of posts and likes, whereas many others post very few. This observation is similar to that of most social media platforms.

**Table 1. Descriptive Statistics**

(a) User-level summary statistics

<table>
<thead>
<tr>
<th></th>
<th>#posts</th>
<th>#Cumulative likes</th>
<th>#Following</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>11.3242</td>
<td>173.1065</td>
<td>31.0892</td>
<td>23.4031</td>
</tr>
<tr>
<td>Std.</td>
<td>16.8707</td>
<td>1476.7562</td>
<td>116.6762</td>
<td>4.7575</td>
</tr>
<tr>
<td>min</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Median</td>
<td>5</td>
<td>25</td>
<td>14</td>
<td>24</td>
</tr>
<tr>
<td>max</td>
<td>561</td>
<td>93376</td>
<td>6130</td>
<td>99</td>
</tr>
</tbody>
</table>

(b) Post-level summary statistics

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std.</th>
<th>Min</th>
<th>Median</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>#likes</td>
<td>25.8841</td>
<td>52.1091</td>
<td>0</td>
<td>10</td>
<td>1747</td>
</tr>
</tbody>
</table>

Note: “#” denotes “the count of.”

**Brand Categorization.** Following Ha et al. (2017), we group the fashion brands into three categories: fast fashion (high street), designer, and mega couture. The categorization is

---

7 We use “fashion bloggers” and “users” interchangeably throughout the paper.
8 The original categorization also includes “small couture” brands, but there is only one observation of such brand in our data. So we consider only three categories.
according to domain experts in the fashion industry, based on brand identity and price ranges. We show some examples of each brand category in Table 2.

\textbf{Table 2. Brand Categorization}

<table>
<thead>
<tr>
<th>Brand Categories</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 1: Fast Fashion (High Street)</td>
<td>Zara, H&amp;M, Forever21</td>
</tr>
<tr>
<td>Level 2: Designer</td>
<td>Kate Spade, Coach, Michael Kors</td>
</tr>
<tr>
<td>Level 3: Mega Couture</td>
<td>Gucci, Prada, Chanel</td>
</tr>
</tbody>
</table>

\subsection*{3.1. Feature Extraction from Images}

For a fashion look, we focus on two key aspects of visual features that can affect one’s utility: the clothing styles and the appearance of the users.\footnote{On the blogging platform, the user himself/herself is the model in the picture. One account consistently posts the account owner’s fashion look.} We describe how we extract and measure the clothing styles in 3.1.1 and user appearance in 3.1.2.

\subsection*{3.1.1. Clothing Styles}

In light of fashion satisfying consumers’ social needs for group cohesion and differentiation (Simmel, 1904) and domain experts’ opinion from well established fashion magazines, at a high level, two style features that are particularly relevant for fashion goods: compatibility and distinctiveness. Compatibility speaks to the combination of clothing items from different categories (e.g., shirts versus pants), whereas distinctiveness measures how visually differentiated each item is from others within the same category. We abstract away from more granular style factors (e.g., color, texture) and capture the styles at a high level, because those granular factors can also be described or evaluated according to compatibility and distinctiveness. Below we explain how we extract these two style features from the fashion looks.
Figure 2. Steps of Clothing Style Features Extraction

To measure fashion styles, we first need to detect or identify the clothing items in each fashion look. We follow the approach of the DeepFashion project by Liu et al. (2016). The method is based on the application of Fast Region-based Convolutional Neural Network (Fast R-CNN) (Girshick, 2015). A Fast R-CNN network takes an image and a set of object location proposals as inputs. It learns to classify objects and refine their spatial locations jointly. We adopt the network architecture of DeepFashion, which was trained on the largest and most comprehensive clothes dataset to date, annotated with clothing landmarks and categories.

For each fashion look, we extract only the clothing items, which are the most visually dominating items in a picture. Most accessories are too small to be precisely detected, so we do not include them in the analysis in this paper. We keep the cropped items (upper and bottom) if the confidence scores are higher than some threshold.\(^\text{10}\) If the detector cannot separate the top and bottom items, we treat the clothes as full-body outfits whose compatibility is assigned an average score of the fashion looks posted during the past three months (same length as a season). After we detect the clothing items, we can proceed to measure compatibility and distinctiveness for the fashion looks.

\(^{10}\) We tried 0.7, 0.8, and 0.9 for robustness checks.
3.1.1.1. Compatibility

Among a large number of fashion looks, we would like to know what clothing items go well together. We adopt a deep learning approach to learn a feature transformation from images of clothing items into a latent space that represents compatibility. We use a Siamese convolutional neural network architecture (Siamese CNN) (Hadsell et al., 2006), where training data are pairs of items that are either incompatible or compatible.

To measure compatibility, we first initialize the model with weights trained on two million pairs of labeled pairs collected from the purchase data of fashion goods on Amazon.com (Veit et al., 2015). As purchasing two items together may not necessarily mean the consumers treat the items as compatible, we further collect an additional training dataset by conducting a survey on Amazon Mechanical Turk (henceforth AMT). We directly ask survey respondents’ opinion on compatibility of randomly selected pairs of items. With transfer learning, we fine-tune the deep neural network with three thousand pairs of responses (compatible versus incompatible) from the survey to improve the measure of compatibility. Please see the appendix for the survey design.

The abstraction of Siamese CNN architecture is shown in Figure 3. Essentially, it learns a feature transformation \( f: I \rightarrow X \) from the image space \( I \) (i.e., raw representation of images) to the style space \( X \) (i.e., another representation that captures the style features). In the style space, compatible items are closer together, and incompatible items are farther away. Then, we can use the distance between two items’ locations in the style space to measure how compatible they are. In Figure 3, \( I_1 \) and \( I_2 \) are the inputs of two clothing items from different categories (top and bottom), \( x_1 \) and \( x_2 \) are vector representations in the style space, \( y \) is the label of data (either compatible or incompatible), and \( \theta \) is the set of parameters that specify the neural network, which we need to estimate.
The loss function $L(\theta)$ is a contrastive loss and can be expressed as:

$$L(\theta) = \sum_{(x_1, x_2)} L_p(x_1, x_2) + \sum_{(x_1, x_3)} L_n(x_1, x_3)$$

The first term $L_p$ penalizes when a compatible pair is too far apart, and the second term $L_n$ penalizes when an incompatible pair is too close compared to some margin.

As illustrated in Figure 4, the embedding of items in the style space are vectors of dimension 256. Following Veit et al. (2015), we measure the compatibility between two items using $L_2$ norm. The architecture of the CNN in Figure 3 is based on one of the most successful network architectures, GoogLeNet (Szegedy et al., 2015), augmented with a 256-dimension fully connected layer.

3.1.1.2. Distinctiveness

We measure the distinctiveness of a clothing item by calculating how visually different the item is from all the other items in the same category. Specifically, we use the embedding in the style space to represent each clothing item’s style and calculate the average style for items posted in the past three months (a season), to account for the fact that one style could be distinctive this season but may not be distinctive later on. The distinctiveness of one item is measured by $L_2$ norm between its style embedding and the average style. This is in a similar spirit to the creativity concept from Toubia and Netzer (2017).
3.1.2. User Appearance

For the model styles, we examine the face and body features, specifically the facial attractiveness and body mass index.

Figure 5. Steps of Model Styles Extraction

3.1.2.1. Facial attractiveness

The first step is to crop the face and get a vector representation of the face. We follow the deep neural network implementation by the Open Face project (Amos et al., 2016). This architecture was trained for face recognition, providing a 128-dimensional intermediate layer that represents a low dimensional embedding of any face image.

The first step is to get the low dimensional features generated using the deep neural network implementation. Then we need to train a supervised learning model to predict attractiveness. Our training data consist of three thousand images with attractiveness scored on a 1 to 7 scale, where
1 means the face is the least attractive and 7 represents the highest value of attractiveness. Each image is labeled by five Amazon mechanical turkers. We take the average of the five ratings for each image as its final rating. Given the continuous nature of the resulting rating, we train a Support Vector Regression (SVR) model that learns the relationship between the 128-dimensional image features and the attractiveness rating. The model achieves high prediction accuracy with a mean absolute error of 0.66 on the test sample.

3.1.2.2. Body feature

We measure BMI to capture the users’ body feature. The training data contains 4206 images of faces with true BMI information, made available by Kocabey et al. (2017). These images are collected from Reddit posts linking to the imgur.com service. With the training data, we first crop the faces and get the embedding of 128 dimensions. Then, with the face embedding as the input and BMI (ranging from 11.5 to 50.8) as the output, we again train an SVR model that learns the relationship between the face image feature and the BMI. Eventually, we can predict BMI for a given fashion look, according to the face detected from the fashion look. The model has good performance with a mean absolute error of 2.45.

3.1.2.3. Gender information enhancement

The gender information for approximately 50% bloggers is not shown on the website. As we also want to know the gender effect, we employ SVM to predict gender from the cropped face images. As a result, we get gender information for 97% bloggers in the whole sample, with 92.89% accuracy on prediction.

3.1.3. Results of Feature Extraction

The objective of the Siamese CNN is to project compatible pairs close together and incompatible pairs far away. Figure 6 plots the distribution of distances for compatible and incompatible pairs for both before and after training for transfer learning. The plots show that the fine-tuned neural network separates the two categories with a greater margin and indicates that the network learned to separate compatible from non-compatible clothing items.
Table 3 reports the performance of the feature extraction tasks on facial attractiveness, BMI, and gender. For SVR task (extracting face attractiveness and BMI), the commonly used performance measure—mean absolute error (MAE)—is reported for the hold-out sample. Accuracy is reported for the binary classification tasks.

In Table 4, we show both high and low score examples of the extracted features, resulting from our trained learning models. Table 5 shows the summary statistics for the style features we extracted from the images.
### Table 4. Example Photos for the Extracted Style Features

<table>
<thead>
<tr>
<th>Feature</th>
<th>Low Score</th>
<th>High Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facial Attractiveness</td>
<td><img src="image1.png" alt="Low Score" /> <img src="image2.png" alt="Low Score" /></td>
<td><img src="image3.png" alt="High Score" /> <img src="image4.png" alt="High Score" /></td>
</tr>
<tr>
<td>Body Feature (BMI)</td>
<td><img src="image5.png" alt="Low Score" /> <img src="image6.png" alt="Low Score" /></td>
<td><img src="image7.png" alt="High Score" /> <img src="image8.png" alt="High Score" /></td>
</tr>
<tr>
<td>Compatibility</td>
<td><img src="image9.png" alt="Low Score" /> <img src="image10.png" alt="Low Score" /></td>
<td><img src="image11.png" alt="High Score" /> <img src="image12.png" alt="High Score" /></td>
</tr>
<tr>
<td>Distinctiveness</td>
<td><img src="image13.png" alt="Low Score" /> <img src="image14.png" alt="Low Score" /></td>
<td><img src="image15.png" alt="High Score" /> <img src="image16.png" alt="High Score" /></td>
</tr>
</tbody>
</table>
Table 5. Summary Statistics for the Extracted Style Features

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facial Attractiveness</td>
<td>4.8201</td>
<td>0.6121</td>
<td>0.3348</td>
<td>7.1819</td>
</tr>
<tr>
<td>Body BMI</td>
<td>27.2535</td>
<td>3.6385</td>
<td>11.4966</td>
<td>50.7738</td>
</tr>
<tr>
<td>Compatibility</td>
<td>47.4943</td>
<td>12.4622</td>
<td>1.7325</td>
<td>257.1844</td>
</tr>
<tr>
<td>Distinctiveness</td>
<td>33.6818</td>
<td>9.8098</td>
<td>8.1391</td>
<td>124.4247</td>
</tr>
</tbody>
</table>

3.2. Exploratory Data Analysis

3.2.1. Users’ Brand and Style Choices

After extracting the style features, we can examine the distribution of styles of fashion looks across the three brand levels to see if the style options are different for each brand level. The boxplots in Figure 7 and the summary statistics in Table 6 show that low-end brands also have pretty good styles, and the style distribution is not much different across brand levels. Therefore, for those who cannot afford luxury brands, there are always substitutes that can provide high-end styles available at lower prices.

Figure 7. Boxplots of Styles across Brands
Table 6. Summary Statistics at Brand-level

<table>
<thead>
<tr>
<th></th>
<th>#item</th>
<th>face</th>
<th>BMI</th>
<th>age</th>
<th>gender</th>
<th>compatibility</th>
<th>distinctiveness</th>
<th>#likes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fast Fashion (High Street) Brands</strong> (Adopted in 99.79% looks)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>1.5652</td>
<td>4.8174</td>
<td>27.2595</td>
<td>24.6837</td>
<td>1.1875</td>
<td>47.4811</td>
<td>33.7177</td>
<td>30.3719</td>
</tr>
<tr>
<td>Std.</td>
<td>0.8540</td>
<td>0.6132</td>
<td>3.6407</td>
<td>5.2284</td>
<td>0.3903</td>
<td>12.4851</td>
<td>9.8417</td>
<td>59.7599</td>
</tr>
<tr>
<td><strong>Designer Brands</strong> (Adopted in 3.05% looks)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>1.0961</td>
<td>4.8533</td>
<td>27.1894</td>
<td>24.8999</td>
<td>1.2192</td>
<td>47.3838</td>
<td>32.9807</td>
<td>57.9064</td>
</tr>
<tr>
<td>Std.</td>
<td>0.2948</td>
<td>0.5515</td>
<td>3.5599</td>
<td>3.4180</td>
<td>0.4139</td>
<td>12.5706</td>
<td>9.1262</td>
<td>106.7643</td>
</tr>
<tr>
<td><strong>Mega Couture Brands</strong> (Adopted in 2.00% looks)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>1.1464</td>
<td>4.9146</td>
<td>26.9429</td>
<td>26.3511</td>
<td>1.1078</td>
<td>48.3523</td>
<td>33.0009</td>
<td>55.9045</td>
</tr>
<tr>
<td>Std.</td>
<td>0.3538</td>
<td>0.6242</td>
<td>3.6141</td>
<td>5.2377</td>
<td>0.3103</td>
<td>11.3491</td>
<td>9.1920</td>
<td>89.4968</td>
</tr>
</tbody>
</table>

We also see from Table 6 that the vast majority of fashion looks contain fast fashion brands. It implies that users tend to use at least one fast fashion item to make the whole ensemble look good. Table 7 shows how users mix and match brands from the three brand categories. For fashion looks that adopt mega couture brands, more than 80% of them also adopt fast fashion brands. For those using designer brands, more than 86% of them mix with fast fashion brands. The data evidence shows the economic significance and importance of consumers’ mixing and matching behavior in fashion choices.

Table 7. Mix-and-match in Fashion Looks Using Higher-end Brands

(a) For looks adopted mega couture brands

<table>
<thead>
<tr>
<th></th>
<th>Only Mega Couture</th>
<th>Mega Couture &amp; Fast Fashion</th>
<th>Mega Couture &amp; Designer</th>
<th>All three</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage</td>
<td>4.6%</td>
<td>80.15%</td>
<td>0.37%</td>
<td>14.89%</td>
</tr>
</tbody>
</table>

(b) For looks adopted designer brands

<table>
<thead>
<tr>
<th></th>
<th>Only Designer</th>
<th>Designer &amp; Fast Fashion</th>
<th>Designer &amp; Mega Couture</th>
<th>All three</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage</td>
<td>3.58%</td>
<td>86.40%</td>
<td>0.24%</td>
<td>9.77%</td>
</tr>
</tbody>
</table>
3.2.2. The Impact of Popularity on Choices and Inter-temporal Tradeoff

To understand how consumers choose brand and style, we need to account for a factor that may strongly affect consumers’ fashion decision in our research context. That is, users are concerned about popularity or peers’ likes. When a user decides whether and what to post, her level of popularity plays a big role. First, being popular can help reduce future cost, given the fact that many popular fashion bloggers are hired by fashion companies and paid to post instead of paying for what they wear. Second, the incentives for posting on social media (Lee et al., 2015) are typically social interaction and self expression, a user may derive higher utility from a given post when she is more popular and has more people watching or following. Therefore, a blogger may be strategic and forward-looking, in the sense that posting to attract others’ likes to build up popularity today can help improve future utility because there would be a larger audience for future posts, as well as lower cost for future posts. In other words, when making a post decision, a blogger considers not only the current period’s utility, but also the future utility. Even though posting may be worse than not posting in the current period, she may still post because it builds up popularity and the blogger can gain much more in the future.

We have explained that the discrete choice of posting is dynamic as popularity affects one’s utility of posting. Moreover, a user also faces inter-temporal tradeoff when making the brand and style choices. Because the brand and style choices affect peer likes a post can attract, but the best choices for attracting likes and building up popularity may not be the choices to satisfy one’s own per-period intrinsic preference. When one is not that popular, she may focus on attracting peer likes and less on self-intrinsic taste. In comparison, a popular blogger can make style and brand choices subject to less financial constraint and focus more on expressing her intrinsic fashion tastes.

An ideal measure of popularity is the number of followers. As we do not observe the number of followers for each user across time, we use the cumulative sum of likes (“SumLike,” henceforth) a person has gotten from previous posts as a good proxy for his/her popularity. Figure 8 shows the positive linear relationship between SumLike and the number of followers at a snapshot time Aug 1st, 2017. The correlation between SumLike and the number of followers by the time Aug 2017 is 0.85, while that between the average number of likes and number of followers is 0.55.
Therefore, we choose SumLike as a proxy measure for one’s popularity or, more generally, the exposure one may get when posting a new fashion look.

*Figure 8. Followers versus Likes*

*Table 8. Popularity Affects Decisions on Fashion Look*

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>DV: Brand 1</th>
<th>DV: Brand 2</th>
<th>DV: Brand 3</th>
<th>DV: Compatibility</th>
<th>DV: Distinctiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Popularity</td>
<td>-0.0115*</td>
<td>-0.00101</td>
<td>0.00278**</td>
<td>0.292***</td>
<td>0.0331</td>
</tr>
<tr>
<td></td>
<td>(0.00609)</td>
<td>(0.00135)</td>
<td>(0.00109)</td>
<td>(0.0628)</td>
<td>(0.0449)</td>
</tr>
<tr>
<td>Observations</td>
<td>18,957</td>
<td>18,957</td>
<td>18,957</td>
<td>18,957</td>
<td>18,957</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.639</td>
<td>0.272</td>
<td>0.391</td>
<td>0.192</td>
<td>0.219</td>
</tr>
</tbody>
</table>

Note: Fixed effects at individual level; Robust standard errors, clustered at individual level.

*p<0.01; **p<0.001; ***p<0.0001

In Table 8, we show the results of five separate regressions to see how popularity up to period \( t - 1 \) affects one’s choice of brands and styles in period \( t \). We can see that when bloggers become more popular, they will post more mega couture brands but fewer fast fashion brands, echoing the cost decreasing effect of being popular. Moreover, they further improve the style, implying that the marginal return of improving style is higher when more people are watching.
3.2.3. Incentives of Posting on Social Media

People post on social media for different reasons: self-expression, social interaction, archiving, escapism, and so on (Lee et al., 2015). Below is the scatter plot of the number of posts and average likes for each user’s posts. We can see that those who post more do not necessarily get more likes than users who post less. Some users’ posts on average attract more than four hundred likes but the users post only around ten times, whereas other users’ posts get few likes but the users post hundreds of times. Therefore, we hypothesize that the users’ utility of posting is not driven only by others’ likes or opinions; they may also derive utility from other channels. For example, an individual might enjoy expressing herself through posting a fashion picture, or she wants to attract those who have the same tastes. In these cases, getting likes is not the primary goal.

As a result, when modeling bloggers’ decision processes, we adopt a general utility functional form, which captures not only the impact of others’ likes but also bloggers’ intrinsic utility derived from the brand and style choices.

*Figure 9. Scatterplot of Posts and Likes*

![Scatterplot of Posts and Likes](image)

However, with the above exploratory analysis, it is still not clear how consumers value brands and styles, and how substitutable brands and styles are. Can good styles from low-end brands compensate for the utility loss from not being able to get the high-end brand with relatively worse styles? We proceed to answer this question by modeling blogger decisions at the micro-level in the next section.
4. Model

The timing of the events is illustrated in Figure 10. In each period:

1. The blogger observes the popularity she has built up with her previous posts (if there are any).
2. She decides whether to post a fashion look. If she posts, she simultaneously chooses the brand and style. If not, she goes to the next period.
3. The number of peer likes for the new post (if there are any) realizes. The next period comes.

*Figure 10. Timing of Events*

4.1. The Basic Model

Let $S_{it}, B_{it}$ denote the decision on style and brand. $P_{it}$ is the binary decision of posting. The per-period utility of blogger $i$ in period $t$ follows a nested constant elasticity of substitution (CES) format (Solow, 1956):
\[ u_{it}(St_{it}, Br_{it}, P_{it} \mid S_{it}, \theta_i) = F_{it}(S_{it}, \theta_i) \cdot R_{it}(St_{it}, Br_{it} \mid \theta_i) \cdot 1(P_{it} = 1) + \epsilon_{it} \] (1)

where \( R_{it}(St_{it}, Br_{it} \mid \theta_i) \) captures the utility gain from the fashion look attributed to the brand \( Br_{it} \) and style \( St_{it} \).

\[ R_{it}(St_{it}, Br_{it} \mid \theta_i) = [\alpha_{i1} Br_{it}^{\rho_{i1}} + (1 - \alpha_{i1}) St_{it}^{\rho_{i1}}]^{1/\rho_{i1}}, \]

where the elasticity of substitution between style and brand is \( \rho_{i1} = \frac{1}{1 - \rho_{it}} \), and \( \alpha_{i1} \) is the share parameter. In our study, we regard the brand value as the incremental value associated with the brand name that is not related to any style attribute. For example, apart from any style features, people choose luxury brands because they work as a social label and provide hedonic rewards.

\( F_{it}(\cdot) \) captures the valuation of a post. It measures the effect of popularity and individual fixed effect on the utility, apart from the brand and style choices.

\[ F_{it}(S_{it}, \theta_i) = \eta_{i0} + \eta_{i1} \ln(1 + SumLike_{i,t-1}). \]

\( SumLike_{i,t-1} \) is the cumulative number of likes user \( i \) got from the fashion looks that she posted in the past \( t - 1 \) periods. In the per-period utility function, \( \epsilon_{it} \) is the brand-and-style choice specific random error, following a Type-1 extreme value distribution.

**Brand choice**

The brand choices are discrete. The blogger makes brand choice \( B_{it} \) from categories of fast fashion (level 1), designer (level 2), and mega couture (level 3), respectively denoted by \( l \in \{1, 2, 3\} \). For each brand level, the blogger decides how many items to include in a post. The brand choice for a fashion post is characterized by the following linear function.\(^{11}\)

\[ Br_{it} = \exp(\sum_{l=1}^{3} \gamma_{l,t} x_{lt,l} + \sum_{l \neq k} \gamma_{l,tk} x_{lt,l} x_{lt,k}), \] (2)

\(^{11}\) We apply \( \exp(\cdot) \) to make sure the brand choice measure is positive so that it is a valid input of the CES function. If there is no brand information for a tagged item, we assume others’ belief is level 1; otherwise, the user would disclose the brand information. This is consistent with the data fact that the majority of brands are level 1 demonstrated in Appendix, and conforms to consumer rationality, given higher-end brands attract more likes.
where \( x_{it,l} \) denotes the number of clothing items of brand category \( l \), in the fashion look posted by blogger \( i \) in period \( t \). \( \gamma_{il} \) captures the utility gain of choosing an additional item of brand \( l \). \( \gamma_{l,ik} \) measures the utility gain of matching items from brand category \( l \) and \( k \).

**Style choice**

The bloggers make style choices. The style, as a factor of the whole fashion look, is incorporated as a sub-nest of the CES utility function.

\[
St_{it} = \left[ \alpha_{i2} f_{it,1}^{\rho_{i2}} + (1 - \alpha_{i2}) f_{it,2}^{\rho_{i2}} \right]^{1/\rho_{i2}},
\]

where \( \alpha_{i2} \) is the share parameter, and \( f_{it,1}, f_{it,2} \) denote the choices of compatibility and distinctiveness. The elasticity of substitution between the two style features is \( \eta_{i2} = \frac{1}{1 - \rho_{i2}} \).

**Budget and cost**

The blogger’s decisions are subject to a budget constraint \( y_i \). Specifically,

\[
\sum_{j=1}^{k} t_{ij} f_{it,j} + \sum_{l=1}^{3} t_{lx,l} x_{it,l} \leq y_i.
\]

The cost of purchasing a clothing item is allowed to change with \( \text{Like}_{l-1} \).

\[
t_{lx,l} = \frac{t_{i,l}}{1 + \delta_l \cdot \ln (1 + \text{SumLike}_{l,t-1})},
\]

where \( l \in \{1,2,3\} \), and \( t_{i,l} \) is the baseline cost for obtaining an item with brand-level \( l \). \( \delta_l \) measures the decreasing effect of the cumulative number of likes on the cost of brand choices. Bloggers may incur less cost if they are valuable to the fashion companies, as influencers can help market their products to the public. The more followers (measured by \( \text{SumLike}_{l,t-1} \)) a user has, the more valuable she is to the fashion company, because her posts will influence a broader
audience who potentially can become customers. This is demonstrated by the fact that many fashion bloggers are paid millions every year by fashion companies.\(^{12}\)

For identification purposes, we normalize the base cost for a fast fashion brand to 1—that is, \(\hat{\ell}_{t,1} = 1\). The budget constraint for blogger \(i\), \(y_t\), measures the highest cost a blogger is willing to pay for the fashion consumption, and it is assumed fixed over weeks, as the time and pecuniary resources allocated to other regular activities (e.g., working, entertainment) typically do not change much over weeks. In estimation, the budget constraint is calculated by the highest ever cost since the user started posting.

### 4.2. State Variables

The state variables are \(S_{it} = \{\text{face}_i, \text{body}_i, \text{age}_i, \text{gender}_i, \text{SumLike}_{i,t-1}, \epsilon_{it}\}\), where \(\text{SumLike}_{i,t-1}\) is time varying while others are fixed for the same user. With \(f_{it} = (f_{it,1}, f_{it,2})\) and \(x_{it} = (x_{it,1}, x_{it,2}, x_{it,3})\), the state transition follows

\[
\text{Like}_{it} = \hat{g}(\text{SumLike}_{i,t-1}, \text{face}_i, \text{body}_i, \text{age}_i, \text{gender}_i, f_{it}, x_{it}) + \zeta_{it}
\]

where \(\hat{g}(\cdot)\) is estimated by a linear regression as the first step, and \(\zeta_{it} \sim N(0, \sigma)\). In our dataset, we observe \(\hat{S}_{it} = \{\text{face}_i, \text{body}_i, \text{age}_i, \text{gender}_i, \text{SumLike}_{i,t-1}\}\). The state transition regression results are shown in section 6.

### 4.3. Inter-temporal Tradeoff

Each user decides on an infinite sequence of decision rules \(\{f_{it}, x_{it}, post_{it}\}_{t=1}^{\infty}\) to maximize the expected discounted utility. Substituting the brand and style choices (equation (2) and (3)) to the per-period utility function (equation (1)), we have

\[
\max_{\{f_{it}, x_{it}, P_{it}\}_{t=0}^{\infty}} E_{\{S_{it}\}_{t=0}^{\infty}} \left\{ \sum_{t=0}^{\infty} \beta^t U_t(f_{it}, x_{it}, P_{it} | S_{it}, \theta_i) | x_{i0}, f_{i0}, P_{i0}, S_{i0} \right\},
\]

where

\(^{12}\)http://www.blogingrace.com/highest-paid-fashion-bloggers/
$$U_t(f_{it}, x_{it}, P_{it} | S_{it}, \theta_i) = 
\eta_{i_0} + \eta_{i_1} \ln(1 + \text{SumLike}_{t-1}) \cdot \left\{ \alpha_i \left[ \exp \left( \sum_{l=1}^{2} \gamma_{l,t} x_{it,l} + \sum_{l \neq k} \gamma_{l,tk} x_{it,l} x_{it,k} \right) \right]^{\rho_{i_1}} + (1 - \alpha_i) \left[ \beta_i f_{i_1}^{\rho_{i_2}} + (1 - \beta_i) f_{i_2}^{\rho_{i_2}} \right]^{\rho_{i_2}} \right\} \cdot 1(P_{it} = 1) + \epsilon_{it}.$$ 

Let $V(S_{it})$ denote the value function:

$$V(S_{it}) = \max_{f_{it}, x_{it}, P_{it}} E_{S_{it}} \{ U_t(f_{it}, x_{it}, P_{it} | S_{it}, \theta_i) + \sum_{\tau = t+1}^{\infty} \beta^{t-\tau} U_{\tau}(f_{it_{\tau}}, x_{it_{\tau}}, P_{it_{\tau}} | S_{it_{\tau}}, \theta_i) | x_{it}, f_{it}, P_{it}, S_{it} \}$$ 

The Bellman Equation (Bellman, 1957) for the dynamic optimization problem is expressed as follows:

$$V(S_{it}) = \max_{f_{it}, x_{it}, P_{it}} E_{S_{it+1}} \{ U_t(f_{it}, x_{it}, P_{it} | S_{it}, \theta_i) + \beta V(S_{it+1}) | x_{it}, f_{it}, P_{it}, S_{it} \}$$

All the decisions, that is, $f_{it}, x_{it}, P_{it}$, are dynamic in the sense that the current period’s decisions affect the next period state through $\text{SumLike}_{t-1}$, which further affects the cost and utility of the user’s brand and style choices. In other words, the users face inter-temporal tradeoffs regarding the utility derived from the post today versus its impact on the utility of posting tomorrow.

### 4.4. Heterogeneity

In the fashion sharing community, users may have different responses to others’ opinions. For example, some users care a lot about others’ opinion (i.e., #likes), whereas some care only about expressing themselves rather than attracting #likes. Similarly, some users base their utility heavily on the brand levels, whereas others care more about the clothing styles. Therefore, we employ a hierarchical Bayesian framework (Rossi et al., 2005) to account for heterogeneity. All structural parameters $\theta_i \in \Theta_i = \{ \alpha_{i_1}, \alpha_{i_2}, \rho_i, \eta_i, t_i, \gamma_i, \delta_i \}$ have random coefficient specification.\footnote{The parameters in bold are vectors with subscription 1,2, etc. as specified in the model.} The prior distribution is normal for $\eta_i, \gamma_i$, and log normal for $\alpha_{i_1}, \alpha_{i_2}, \rho_i, t_i, \gamma_i, \delta_i$, where $\alpha_i = \frac{1}{\gamma_i}$. The prior distribution is specified as $N(\mu_\theta, \sigma^2_\theta)$. We use diffuse hyper-prior distribution for all parameters.
5. Identification and Estimation

5.1. Identification

In our model, the unknown parameters include those in the state transition process \( (\tilde{b}) \) and the primitives in the utility function \( (\Theta_t) \). We briefly explain the identification of the key parameters.

First of all, the transition function parameters can be identified by the variation in the number of likes \( \text{Like}_{l,t} \) and the corresponding fashion brands, style features, and blogger demographics (i.e., \( \text{SumLike}_{l,t-1}, \text{face}_l, \text{body}_l, \text{age}_l, \text{gender}_l, f_{it}, x_{it} \)). For the discount factor \( \beta \), as acknowledged in the literature (e.g., Rust, 1987), it cannot be separately identified from the utility parameter, that is, the individual-fixed effect in our context. Therefore, following the conventional approach, we fixed the weekly discount at 0.998 to stay consistent with the literature (Hartmann & Nair, 2010; Liu et al., 2018).

Table 9. Summary of the Parameters

<table>
<thead>
<tr>
<th>Notation</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha_{l1} )</td>
<td>Share parameter for brand and style.</td>
</tr>
<tr>
<td>( \alpha_{l2} )</td>
<td>Share parameter for different style features.</td>
</tr>
<tr>
<td>( \rho_l )</td>
<td>Elasticity of substitution, ( \rho_l = {\rho_{l1}, \rho_{l2}} ), for brand versus style, and between style features.</td>
</tr>
<tr>
<td>( \eta_l )</td>
<td>Governing the efficiency or productivity. ( \eta_l = {\eta_{l0}, \eta_{l1}} ), where ( \eta_{l0} ) captures individual fixed-effect, ( \eta_{l1} ) measures the effect from the cumulative likes.</td>
</tr>
<tr>
<td>( t_l )</td>
<td>Cost parameters for brand and style choices. ( t_l = {t_{l1}, t_{l2}, t_{lx2}, t_{lx3}} ).</td>
</tr>
<tr>
<td>( \gamma_l )</td>
<td>Utility gain from an item of a certain brand level. ( \gamma_l = {\gamma_{l1}, \gamma_{l2}, \gamma_{l3}, \gamma_{l1,2}, \gamma_{l1,3}, \gamma_{l2,3}} ).</td>
</tr>
<tr>
<td>( \delta_l )</td>
<td>The costing decreasing effect from cumulative #likes (as a proxy for #followers).</td>
</tr>
</tbody>
</table>

The elasticity of substitution between brand and style can be identified by the variation in the brand \( (B\gamma_{l,t}) \) and style choices \( (St_{t,t}) \) across time. Similarly, the elasticity of substitution between compatibility and distinctiveness can be identified by the variation in the two style choices across periods. The brand and style choices help us back out the underlying cost of each blogger. We
can separately identify the effect of $\text{SumLike}$ on the utility ($\eta_{i1}$) and on the cost ($\delta_i$), because the cost decreasing effect is only on the brand choices, whereas $\eta_{i1}$ affects both the brand and style choices in the same way. So, the different evolving patterns of brand and style choices can help us identify $\eta_{i1}$ and $\delta_i$.

In summary, the structural parameters to be estimated are $\Theta_i = \{\alpha_{i1}, \alpha_{i2}, \rho_i, \eta_i, t_i, y_i, \delta_i\}$.

5.2. Likelihood

The likelihood function is

$$
\text{Likelihood} = L \left( \left\{ \left\{ f_{it}^*, x_{it}, P_{it} \right\}_{i=1}^{T} \right\}_{t=1}^{I} \right) L \left( \left\{ S_{it} \mid S_{it-1}, f_{it}, x_{it}, P_{it} \right\}_{i=1}^{I} \right) L \left( \left\{ S_{i0} \right\}_{i=1}^{I} \right)
$$

where $S_{it}$ includes all the observable states. According to the above likelihood function, the likelihood for the state transition process and that for the optimal choices can be separately estimated. Our data cover the entire history of activities for each individual, and everyone starts with $\text{Like}_{i0} = 0$. As the first step, we estimate the state transition process, $\mathcal{G}(\cdot)$, with a linear regression. Then we maximize the likelihood of the optimal choices:

$$
L \left( \left\{ \left\{ f_{it}^*, x_{it}, P_{it} \right\}_{i=1}^{T} \right\}_{t=1}^{I} \right) = \prod_{t=1}^{T} \prod_{i=1}^{I} L \left( \left\{ f_{it}, x_{it}, P_{it} \mid S_{it}, \theta_i \right\} \right),
$$

where the brand choices $x_{it}$ are discrete and style choices $f_{it}$ are continuous.

The likelihood for each choice $\{f_{it}, x_{it}, P_{it}\}$, consisting of both discrete and continuous choices, can be calculated through a discrete way. Note that for each combination of choice, $\{f_{it}, x_{it}, P_{it}\}$, the predicted number of likes can be obtained with the estimated transition regression model—that is, $\text{Like}_{it} = \mathcal{G}(\text{Like}_{i,t-1}, face_i, body_i, age_i, gender_i, f_{it}, x_{it}) \cdot 1\{P_{it} = 1\}$. Then, we can first eliminate all style choices that are strictly dominated. Specifically, the assumption of blogger rationality implies a unique choice set $\{f_{it}, x_{it}, P_{it}\}$ corresponding to the set $\{\text{Like}_{it}, x_{it}, P_{it}\}$. With the CES utility functional form, there exist unique closed-form solutions for $f_{it}$ given $\{\text{Like}_{it}, x_{it}, P_{it}\}$. To put it in math, given $\{x_{it}, P_{it}\}$, the optimal choices of $f_{it}$ are unique for each $\text{Like}_{it}$, obtained by solving the following maximization problem:

$$
\max_{f_{ijt}} \left[ \beta_i f_{1it}^{\rho_{i1}} + (1 - \beta_i) f_{2it}^{\rho_{i2}} \right]^{1/\rho_{i2}}
$$
subject to $\hat{g}\{f_{it} | x_{it}, P_{it}, S_{lt}\} = \text{Like}_{lt}$ and $\Sigma_{j=1}^{k} t_j f_{ijt} + \Sigma_{l=1}^{3} t_{xl} x_{ltl} \leq y_i$.

Therefore, styles choices that satisfy $f_{it} \neq f_{it}^*$ but lead to the same $\text{Like}_{lt}$ are strictly dominated, thus will never be chosen. With this observation, we can further relieve the computation burden by acting on $\text{Like}_{lt}$ instead of multiple choices $f_{it}$. Please see the appendix for details about how we transform the continuous choice space for $f_{it}$ into a discrete action space on $\text{Like}_{lt}$ while preserving the continuous nature of the style choices.

The likelihood of the optimal choices can be expressed as

$$L\left(\{f_{it}, x_{it}, P_{it} | S_{lt}, \theta_t\}_{t=1}^{T}\right) = \prod_{t=1}^{T} \prod_{l=1}^{L} L\{f_{it}, x_{it}, P_{it} | S_{lt}, \theta_t\} = \prod_{t=1}^{T} \prod_{l=1}^{L} \mathbb{P}_{\text{Like}_{lt}, x_{it}, P_{it} | S_{lt}, \theta_t}.$$  

With Type-1 extreme value distribution for the random error, the choice probability can be written as:

$$\mathbb{P}_{\text{Like}_{lt,n}, x_{it,n}, P_{it,n} | S_{lt}, \theta_t} = \frac{\exp\{v(\text{Like}_{lt,n}, x_{it,n}, P_{it,n} | S_{lt}, \theta_t)\}}{\sum_{n=1}^{N} \exp\{v(\text{Like}_{lt,n}, x_{it,n}, P_{it,n} | S_{lt}, \theta_t)\}}$$

where

$$v(\text{Like}_{lt,n}, x_{it,n}, P_{it,n} | S_{lt}) = U_t \left(f_{it,n}, x_{it,n}, P_{it,n} | S_{lt}\right) - \epsilon_{it,n},$$

and $f_{it,n}$ are observed from the data or backed out from $\{\text{Like}_{lt,n}, x_{it,n}\}$ if there is no post in that period. $N$ is the total number of discrete choices. According to the data, in a fashion look, there are up to three items tagged as brand level 1, while there are up to two items tagged as brand level 2 (designer) and 3 (couture). For fashion looks without brand tags, we assume audiences’ belief about the brand level is fast fashion (level 1), which is consistent with the majority observation and rational behavior of tagging. Therefore, there are $3 \times 2 \times 2 - 1 = 35$ brand choices if one decides to post in our context. The choice of target #likes can be from 0 to the largest #likes the bloggers have gotten on the website, which is 1747. Together with the choice

---

14 There are only fewer than 1% posts tagging more than the upper limits of brands; therefore, we bound the choices below 3, 2, 2 respectively for each level of brand. Though there are cases where people wear the same clothing item in multiple posts, they do not tag the same item, and there must some other new items across different posts.

15 If the items are of higher-level brands, the consumer would tag them given the belief that no-tagged brands are level 1, because the fixed-effect regression indicates that a higher-level brand contributes positively to #likes.
of not posting, the bloggers have in total $N = 35M + 1$ discrete choices. For ease of computation, we can shrink the choice space for an individual according to the historical #likes she has gotten, plus some deviation based on $\sigma$.

5.3. Estimation Methods

To estimate the infinite horizon dynamic structural model, we explore several methods, including the conditional choice probability estimation (CCP) (Hotz & Miller, 1993; Aguirregabiria & Mira, 2007; Arcidiacono & Miller, 2011), the simulated method of moments (SMM) (Pakes & Pollard, 1989; McFadden, 1989), and the Bayesian estimation method (Imai, Jain, & Ching, 2009 (IJC)). SMM matches data moments with simulated moments, but it requires fully solving the dynamic optimization problem and is thus computationally very costly; moreover, it cannot capture the rich heterogeneous responses across different individuals. The CCP methods improve on the computational efficiency but can recover only very limited heterogeneity. The IJC method serves our goal to capture rich individual responses, and the computational burden is also alleviated because it requires evaluating the value function only once in each iteration.

However, IJC is designed for discrete choice models, but our model includes both discrete and continuous choices. With the modification for the dynamic choice problem explained in section 5.2 (please see more details in Appendix D), we can apply the IJC while reserving the continuous nature of the style choices.

6. Results

6.1. Model Comparison

We compare our model with the benchmark model without the forward looking and direct utility derived from followers. When bloggers are myopic, they do not consider the impact of today’s choice on tomorrow’s state, that is, number of followers. When there is no direct utility gain from followers, the only effect of followers is through decreasing cost. The utility gain from followers adds more to the model fitting than forward looking. All four measures show that the proposed model outperforms the benchmarks significantly.
Table 10. The Proposed Model vs. Alternative Models

<table>
<thead>
<tr>
<th></th>
<th>No forward-looking</th>
<th>No utility from followers</th>
<th>Proposed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hit Rate: Post</td>
<td>0.564</td>
<td>0.532</td>
<td>0.877</td>
</tr>
<tr>
<td>Hit Rate: Fast Fashion</td>
<td>0.775</td>
<td>0.610</td>
<td>0.922</td>
</tr>
<tr>
<td>Hit Rate: Designer</td>
<td>0.589</td>
<td>0.533</td>
<td>0.897</td>
</tr>
<tr>
<td>Hit Rate: Mega Couture</td>
<td>0.520</td>
<td>0.421</td>
<td>0.811</td>
</tr>
</tbody>
</table>

6.2. Parameter Estimates

Table 11 shows the transition process estimated with an OLS regression. We can see that popularity does have a substantial positive effect on the peer likes for a new fashion post. The number of clothing items has increasingly positive effects with higher brand levels. More attractive faces attract more likes. The average audience likes a lower BMI, younger looks, and male models. We also see a significant positive effect of compatibility. The distinctiveness does not show a significant impact on attracting likes. The negative effect of the interaction term between high-street brands and mega couture brands shows that people do not respond favorably to outfits that match low and high-end brands together. However, the positive effect of including either a fast fashion brand or a mega couture brand dominates the negative interaction effect.

Table 12 reports the results for structural parameters. The estimation converges with 7000 iterations. We ran 10,000 iterations and report the results using the last 2000 iterations after burn-in. Figure 11 shows the histogram of some structural parameters across individual bloggers. The mean elasticity of substitution is $r_{1l} = 1/(1 - \rho_{1l}) = 1.24$ and implies that styles and brands are quite substitutable for most bloggers. In other words, they find it easy to substitute a high-end brand with good style to derive the same utility. As shown in Figure 11, 60.11% of the individuals treat style and brands as substitutes. For this group of consumers, high-end brands may need to worry about the copycats’ cannibalization effect. For the rest 39.89% users, they view style and brand as complements, meaning they will not be lured by only good styles, which makes the copycat problem less worrisome.

---

16 Convergence was visually assessed with plots of the structural parameters. We store 100 past pseudo-value functions.
Table 11. Regression of State Transition

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>OLS</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log (1+#Like_{t-1})</td>
<td>12.74***</td>
<td>0.147</td>
</tr>
<tr>
<td># Fast fashion (level 1)</td>
<td>8.814***</td>
<td>0.466</td>
</tr>
<tr>
<td># Designer (level 2)</td>
<td>15.72**</td>
<td>7.237</td>
</tr>
<tr>
<td># Mega couture (level 3)</td>
<td>24.18***</td>
<td>7.180</td>
</tr>
<tr>
<td>Face</td>
<td>5.768***</td>
<td>0.744</td>
</tr>
<tr>
<td>BMI</td>
<td>-0.644***</td>
<td>0.121</td>
</tr>
<tr>
<td>Age</td>
<td>-0.485***</td>
<td>0.0819</td>
</tr>
<tr>
<td>Gender</td>
<td>2.952***</td>
<td>0.928</td>
</tr>
<tr>
<td>Compatibility</td>
<td>0.151***</td>
<td>0.0351</td>
</tr>
<tr>
<td>Distinctiveness</td>
<td>0.0116</td>
<td>0.0448</td>
</tr>
<tr>
<td>Level 1*Level 2</td>
<td>-0.604</td>
<td>2.812</td>
</tr>
<tr>
<td>Level 1*Level 3</td>
<td>-6.937**</td>
<td>2.963</td>
</tr>
<tr>
<td>Level 2*Level 3</td>
<td>-1.087</td>
<td>6.935</td>
</tr>
<tr>
<td>Constant</td>
<td>-45.87***</td>
<td>7.532</td>
</tr>
</tbody>
</table>

| Observations                     | 21,093   |
| R-squared                        | 0.304    |

Note: "#" denotes "the count of"; Robust standard errors; Observations are collapsed at weekly level; *p<0.01, **p<0.001, ***p<0.0001.

The estimates of $\eta_{14}$ indicate that most people (about 80%) value popularity or others’ attention, but there exists large heterogeneity across individuals, with a standard deviation of 0.3258. There are about 21.92% users have negative value for popularity, probably because they feel uncomfortable or stressful when more people are watching and paying attention to them. This group of users may post for archiving or escapism, which requires less or no other people’s attention. From the third plot of Figure 11, we see there is also a large cost decreasing effect of popularity. This is consistent with the fact that influential bloggers are hired by fashion companies to promote their products.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Interpretation</th>
<th>Mean ($\mu_\theta$)</th>
<th>Standard deviation ($\sigma_\theta$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_{i1}$</td>
<td>Share parameter</td>
<td>0.5143 (0.0059)</td>
<td>0.1068 (0.0181)</td>
</tr>
<tr>
<td>$\alpha_{i2}$</td>
<td></td>
<td>0.5812 (0.0054)</td>
<td>0.0971 (0.0172)</td>
</tr>
<tr>
<td>$\rho_{1i}$</td>
<td>Elasticity of substitution</td>
<td>0.1903 (0.0173)</td>
<td>0.2759 (0.0135)</td>
</tr>
<tr>
<td>$\rho_{2i}$</td>
<td></td>
<td>-0.1269 (0.0221)</td>
<td>0.4131 (0.0151)</td>
</tr>
<tr>
<td>$\eta_{0i}$</td>
<td>Fixed effect</td>
<td>0.1557 (0.0265)</td>
<td>0.4497 (0.0205)</td>
</tr>
<tr>
<td>$\eta_{1i}$</td>
<td>Effect of cumulative #likes</td>
<td>0.4926 (0.0181)</td>
<td>0.3258 (0.0152)</td>
</tr>
<tr>
<td>$t_{1i}$</td>
<td>Cost for style choices</td>
<td>1.0768 (0.0263)</td>
<td>0.4277 (0.0178)</td>
</tr>
<tr>
<td>$t_{2i}$</td>
<td></td>
<td>0.9369 (0.0213)</td>
<td>0.3713 (0.0172)</td>
</tr>
<tr>
<td>$t_{x2,i}$</td>
<td>Cost for brand choices</td>
<td>2.0601 (0.0224)</td>
<td>0.4103 (0.0143)</td>
</tr>
<tr>
<td>$t_{x3,i}$</td>
<td></td>
<td>2.9698 (0.0292)</td>
<td>0.8729 (0.0177)</td>
</tr>
<tr>
<td>$\gamma_{1i}$</td>
<td></td>
<td>1.1805 (0.0311)</td>
<td>0.4763 (0.0174)</td>
</tr>
<tr>
<td>$\gamma_{2i}$</td>
<td></td>
<td>0.6857 (0.0161)</td>
<td>0.2827 (0.0162)</td>
</tr>
<tr>
<td>$\gamma_{3i}$</td>
<td>Gain from brand choices</td>
<td>0.7848 (0.0208)</td>
<td>0.2917 (0.0178)</td>
</tr>
<tr>
<td>$\gamma_{12,i}$</td>
<td></td>
<td>-0.0130 (0.0208)</td>
<td>0.3994 (0.0178)</td>
</tr>
<tr>
<td>$\gamma_{13,i}$</td>
<td></td>
<td>-0.0269 (0.0228)</td>
<td>0.4013 (0.0176)</td>
</tr>
<tr>
<td>$\gamma_{23,i}$</td>
<td></td>
<td>0.0050 (0.0209)</td>
<td>0.3757 (0.0146)</td>
</tr>
<tr>
<td>$\delta_i$</td>
<td>Cost decreasing effect from cumulative #likes</td>
<td>0.9902 (0.0220)</td>
<td>0.3821 (0.0156)</td>
</tr>
</tbody>
</table>

Note: Standard errors in parentheses.
Figure 11. Distribution of Structural Parameters across Individuals

Elasticity of Substitution between Brand and Style

- Substitutes: 60.11%
- Complements: 39.89%

Histogram of Followers' Effect on Utility

- 78.08%
- 21.92%

Histogram of Followers' Cost Effect

Cost Decreasing Effect
7. Counterfactual Studies

7.1. If Fast Fashion Cannot Copy Mega Couture Styles

We wanted to see what would happen to consumers’ choices if copyright law provided more protection for fashion designs. In this counterfactual world, the fast fashion companies are prohibited from producing styles similar to the original fashion styles of mega couture brands. As a result, the style choices at the fast fashion level would be more restricted. Specifically, the styles of the high-end brands become relatively more exclusive or distinctive, whereas the styles at the fast fashion level become not distinctive, according to our measurement of distinctiveness. Moreover, as fast fashion companies cannot produce copycat styles in the counterfactual world, if consumers wear both the high-end and low-end styles together, the compatibility would be much lower than before. How would exactly the style options change? For each mega couture item, we first calculate its distance to all other fast fashion items, using the vector representation. We experimented by dropping the fast fashion styles with 10% and 5% smallest distance (largest similarity) and replace the removed ones with the average of the other fast fashion styles. Then we calculate the distinctiveness for the fast fashion styles, and the highest compatibility between a fast fashion and a higher-end brand. We observe the remaining fast fashion styles are below about 25 percentile of the mega couture styles. Therefore, we operationalize the analysis by restricting the distinctiveness of the fast-fashion styles to be bounded below 25 percentiles of the styles from the mega couture brands. We also restrict the compatibility of matching fast fashion with mega couture brands to be within the bottom quartile of other combinations.

The first column of Table 13 shows the results of a hundred simulations. On average, the posting probability drops by 7.92%. Interestingly, not only the fast fashion brands are worse off, but also the adoption of all brands decreases. Why are the high-end brands also worse off when copycats are prohibited? By comparing consumers’ choices in the counterfactual world and what they chose before, we found three mechanisms contributing to a lower demand for the mega couture brands.

First, many bloggers combine clothing items across brand levels to make a complete outfit. When the style choice is restricted at low-end brands, consumers are subject to higher financial pressure to buy high-end brands to get a satisfactory ensemble. In other words, this is driven by
the consumers who cannot always afford high-end brands for their ensembles and who therefore mix and match both low-end and high-end brands. The counterfactual policy would put this group of consumers unsatisfied with the styles of mixing and matching high-end and low-end brands together, resulting in them buying nothing. Thus, they end up buying fewer clothes and post less. The data facts (Table 7) suggest that about 80% (86%) of fashion looks incorporating a mega couture (designer) brand also include items at the fast-fashion level, further demonstrating the economic significance of the mix-and-match mechanism in the fashion market.

Second, some consumers may not value high-end brands that much. But they can accumulate popularity across time by wearing attractive styles from fast fashion. Once their popularity is high, there are more people following and paying attention to them, then they will derive more value from what they wear, including high-end brands. Therefore, they will be more likely to adopt high-end brands.

Third, many consumers cannot afford mega couture brands in the very beginning, so they rely on trendy styles from the low-end brands to build up popularity. Once the popularity is high enough, they start to be able to afford more high-end brands due to the cost decreasing effect of popularity. The three above-mentioned mechanisms are market expansion effect brought by fashion copycats. Our findings demonstrate that copycats’ market expansion effect dominates the competition effect, leading to a positive net effect on the demand of high-end brands. In addition, the results also show a boost in the choice of compatibility, which for many bloggers is a substitute for distinctiveness.

### 7.1.1. Discussion on Firm Strategy

As we do not have firm side data and only model the consumers’ decision making, our results speak to the cases where fast fashion firms do not find a way to dramatically change their styles and firms charge the original price. This section discusses how firms would react to the counterfactual policy and the corresponding demand effect on high-end brand.

Fast fashion firms would either put the same effort in designing their own styles, which is less distinctive and matches not well with higher-end brands, but decrease the price to attract more consumers. This escalated price competition would further decrease the demand for mega couture brands. Alternatively, the low-end brands may invest more in coming up a large number
of their own styles to make them not only distinctive but also compatible with higher-end brands. We keep ignorant about how they can achieve this goal with the current low price, but one possibility is that they may transform into designer brands, providing better styles but also higher price due to the higher production cost. The demand effect of prohibiting copycats on mega couture brands is most likely still negative because there is only a small proportion of consumers mix and match designer brands with mega couture brands.

Regarding mega couture brands’ strategic reaction, they may start a low-end product line, providing similar styles to those from high-end product line but charge lower price (i.e., umbrella branding). However, in this case, the low-end product line may erode the parent brand’s value, and the demand effect is not clear without additional information and further empirical study. The other strategy the mega couture brands may use is to simply lower the price thus attract more demand. In this case, the mega couture brands’ profit is definitely lower than before because they have to charge a lower price to achieve the same demand.

Overall, we can see if fast fashion copycats were prohibited, neither fast fashion firms nor mega couture brands can find an easy way to combat the loss.

7.1.2. Discussion on Generalizability

As we acknowledged upfront, an average consumer may differ from a fashion social media user. To generalize to all other consumers who do not use social media or do not post fashion content on social media, we need further study and additional data which is not available for now. However, we still can try to make the results more generalizable. In our sample, there are around 10% users are professional fashion bloggers who operate their own blogging website. They may have very different incentive than an average consumer. So we exclude these professional bloggers and check the choice change for other users. The results are reported in the second column of Table 13. Compared to the first column, we see the net effect is still negative but with smaller magnitude. This is due to the fact that professional bloggers care more about popularity, and the lack of fast fashion copycats makes it harder to build up popularity. As a result, the value increasing effect and cost decreasing effect of popularity cannot work their way to increase adoption of mega couture brands.
7.2. Offline Market: Absence of Follower Effect

This provides a similar scenario as the offline fashion market, where people do not have the “likes” information for what they wear. The posting decision is analogous to the purchase decision. There are two consequences. First, one’s utility does not involve the likes’ impact; second, the cost of purchasing will not decrease over time.

We seek to evaluate how less/more likely a blogger would post. This analysis allows us to understand how many more purchases can be achieved by the existence of social media, compared to the traditional offline market. Companies can leverage the “follower function” of social media to boost their revenue. The results in the fourth column of Table 13 show that there would be a 9.07% drop in post probability if bloggers did not know about others’ following. High-end brands suffer most with the largest drop in choice probability, because the cost decreasing effect is also deprived, and this affects the purchase of high-end brands most.

7.3. When Brands Are Not Observable

As previously mentioned, our data context has a special brand tagging function, which requires users to pinpoint the brands for the clothing items in their fashion look and clearly lists the brands besides the picture. In most other social medium, there is no such feature. Moreover, in the offline market, brand logos are usually hidden or not obvious. In this counterfactual study, we investigate how fashion bloggers make choices if there were no way to inform others about the brand information. Consistent with the main analysis, when there is no brand information for a clothing item, the belief is that it is a high street brand (i.e., low-end brand). The results are shown in the third column of Table 13.

The results show that there would be 0.21% fewer posts. The users would post 0.74% more fast fashion brands, but 0.25% fewer designer brands and 0.30% fewer mega couture brands. Moreover, both style features are improved. Intuitively, when higher-end brands cannot be identified, bloggers cannot use high-end brands to attract peer likes anymore. Therefore, for those who care about popularity and peer likes a lot, they have fewer means to achieve the goal, resulting in lower utility from blogging and thus fewer posts. On the other hand, they switch focus to either using more low-end brands or investing more on styles. As the cost of a low-end brand is lower, bloggers would be able to purchase more low-end clothes, resulting in more posts.
of fast fashion brands. In summary, for higher-end brands, hiding or obscuring the brand logo would benefit the lower-end brands but hurt high-end brands, given all brands can provide similar styles. This result implies that, high-end brands may want to design their brand logos obvious enough to consumers, so that those who care about peer likes would have greater incentive to buy.

7.4. **Platform Ranking System: If “Sensitive” Users Are Prioritized**

As bloggers value others’ opinions differently, the website can change the ranking system so that it favors those “sensitive” bloggers, to incentivize more fashion posts. Effectively, the platform can exogenously change the number of likes a post can get through the ranking system. We run a counterfactual analysis on the top 10% of sensitive bloggers by exogenously giving one more like to their posts. The results in the last column of Table 13 show that, on average, there would be a 0.90% increase in the probability of posting in each period. There is a 0.61%, 0.87%, and 0.89% increase in the number of items posted from brand levels 1, 2, and 3 respectively.

*Table 13. Results of Counterfactual Studies*

<table>
<thead>
<tr>
<th>Counterfactual Policies</th>
<th>No copycats</th>
<th>No copycats (Sub sample)</th>
<th>No brand info</th>
<th>No followers</th>
<th>Targeted ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>Δ#Posts (%)</td>
<td>-7.92%</td>
<td>-6.12%</td>
<td>-0.21%</td>
<td>-9.07%</td>
<td>0.90%</td>
</tr>
<tr>
<td>Δ#Fast fashion (%)</td>
<td>-5.78%</td>
<td>-5.52%</td>
<td>0.74%</td>
<td>-6.07%</td>
<td>0.61%</td>
</tr>
<tr>
<td>Δ#Designer (%)</td>
<td>-7.68%</td>
<td>-6.08%</td>
<td>-0.25%</td>
<td>-11.78%</td>
<td>0.87%</td>
</tr>
<tr>
<td>Δ#Mega couture (%)</td>
<td>-12.44%</td>
<td>-9.32%</td>
<td>-0.30%</td>
<td>-19.86%</td>
<td>0.89%</td>
</tr>
<tr>
<td>ΔCompatibility (%)</td>
<td>-3.34%</td>
<td>-3.51%</td>
<td>99.93%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ΔDistinctiveness (%)</td>
<td>-23.87%</td>
<td>-15.76%</td>
<td>19.88%</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

8. **Contribution and Limitations**

In the fashion market, it has been argued that the fast fashion brands copy the designs of the high-end fashion brands. This practice can potentially reduce the distinctiveness of the luxury
fashion brands thus erode their brand equity. However, there is no systematic study attempting to investigate the impact of fast fashion copycats on high-end brand equity and the underlying reasons. The key challenges limiting such study are the lack of scalable ways to quantify fashion styles and the unavailability of large-scale data on individuals’ choices over brands and styles. In this paper, we use the user-generated data from a large fashion sharing platform and the state-of-the-art deep learning methods on image analytics to quantify fashion styles. Given fashion social media users’ significant impact on fashion trend and demand, understanding their decision process sheds light on fashion consumers’ choices across population. For fashion goods, brand and style are two of the product attributes consumers care about most. Incentive-wise, fashion bloggers make the tradeoff between self-intrinsic tastes and building up popularity. To figure out the underlying reasons of how fast fashion may affect high-end brands’ demand, we build a structural model to investigate fashion social media users’ decision processes that reveals their heterogeneous responses to brands, styles, and popularity.

Our results show that styles and brands are quite substitutable for most people. In other words, they find it easy to substitute a high-end brand with good styles to derive the same utility. These are the consumers that high-end brands could lose to the low-end brands providing comparable styles. We also find that most users value being popular (or peer likes), but there exists a large variance in how much they value popularity. This variation explains why some people keep posting even though they get almost zero peer likes all the time. Moreover, we find that a higher popularity can help reduce posting cost which is consistent with the fact that fashion bloggers with lots of followers are sponsored by fashion companies to post about their products.

In the main counterfactual analysis, we restrict the availability of style choice for fast fashion brands. The results show that not only the fast fashion brands will suffer, but also the high-end brands will be worse off. This means that a more restrictive copyright law on fashion design may not necessarily help the mega couture brands. We found three mechanisms that contribute to this result. Because the more restricted choice of styles from low-end brands would limit the mix-and-match choices for consumers and put them on greater financial constraint to get a satisfactory ensemble of clothes, resulting in them buying less high-end brands. Moreover, the lack of good styles from low-end brands will make it harder for consumers to build up popularity/likeability, which limits the value of what they wear, including high-end brands. A low popularity also makes it unlikely to get a cost reduction from high-end brands. All these
reasons indicate that copycats can benefit the high-end brand, as our findings suggest the market expansion effect dominates the competition effect. We also simulate the case where brand information cannot be seen, which is similar to the offline market where brand logos are hidden or other social media where brand tagging is not featured. The results indicate that on average, fast fashion benefits, whereas the high-end brands suffer. Another counterfactual analysis indicates that there would be a 9.07% drop in post probability if users did not know about peer likes, and high-end brands would suffer the most. This scenario is similar to that of the offline market. This analysis, therefore, allows us to understand how many more purchases can be achieved by the mere existence of social media, compared to the traditional offline market. We also find that an alternative ranking system that prioritizes people who more highly value “likes” can hugely increase the amount of content generated on the platform.

There are several limitations of the current paper that call for future studies. First, the style measure is an objective measure, whereas different consumers may evaluate the compatibility and distinctiveness differently. So the interpretation of the results should be based on the objective style measure we use. An ideal scenario is that we get to know how each consumer evaluates styles. Though this is not feasible, future research could do a more targeted analysis if there were more granular information about consumers from lab experiments or surveys and could then match people’s evaluations of styles based on these observables. Second, though we tried to capture the most salient factors in a fashion post—that is, the model face, the body feature, and the style of clothing items—we did not include everything that may affect the likes a post can get. For example, the accessories may matter. The challenge is that although the detection algorithms for fashion items are state-of-the-art, they are not perfect, especially for items that are too small (e.g., earrings and hats are comparatively quite small). Had we been able to use a better detection algorithm or rich training set, we could have incorporated more factors in the model.
9. Reference


Appendix

A. More on summary statistics

Table 14 shows how many times each brand appears in the posts in total.

*Table 14. Frequency of brands (Top 60)*

<table>
<thead>
<tr>
<th>Brand</th>
<th>Count</th>
<th>Brand</th>
<th>Count</th>
<th>Brand</th>
<th>Count</th>
<th>Brand</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>H&amp;M</td>
<td>35891</td>
<td>Dr. Martens</td>
<td>2743</td>
<td>GUESS</td>
<td>1490</td>
<td>Celine</td>
<td>895</td>
</tr>
<tr>
<td>Zara</td>
<td>28865</td>
<td>Topman</td>
<td>2546</td>
<td>Adidas</td>
<td>1393</td>
<td>Prada</td>
<td>876</td>
</tr>
<tr>
<td>Forever 21</td>
<td>15384</td>
<td>Stradivarius</td>
<td>2376</td>
<td>Diy</td>
<td>1382</td>
<td>Choies</td>
<td>847</td>
</tr>
<tr>
<td>Topshop</td>
<td>10273</td>
<td>River Island</td>
<td>2325</td>
<td>Cheap Monday</td>
<td>1359</td>
<td>Blanco</td>
<td>793</td>
</tr>
<tr>
<td>Vintage</td>
<td>8493</td>
<td>Aldo</td>
<td>2265</td>
<td>Shein</td>
<td>1306</td>
<td>Gucci</td>
<td>788</td>
</tr>
<tr>
<td>Mango</td>
<td>5518</td>
<td>Pull&amp;Bear</td>
<td>2169</td>
<td>Uniqlo</td>
<td>1288</td>
<td>J. Crew</td>
<td>783</td>
</tr>
<tr>
<td>Primark</td>
<td>5504</td>
<td>Nike</td>
<td>2160</td>
<td>Marc by Marc Jacobs</td>
<td>1258</td>
<td>Ralph Lauren</td>
<td>776</td>
</tr>
<tr>
<td>Asos</td>
<td>5329</td>
<td>Thrifted</td>
<td>2051</td>
<td>Louis Vuitton</td>
<td>1156</td>
<td>C&amp;A</td>
<td>772</td>
</tr>
<tr>
<td>American Apparel</td>
<td>5006</td>
<td>Jeffrey Campbell</td>
<td>2040</td>
<td>Cotton On</td>
<td>1139</td>
<td>Vero Moda</td>
<td>768</td>
</tr>
<tr>
<td>Converse</td>
<td>3922</td>
<td>Vans</td>
<td>1978</td>
<td>New Yorker</td>
<td>1106</td>
<td>Old Navy</td>
<td>755</td>
</tr>
<tr>
<td>Ray-Ban</td>
<td>3716</td>
<td>Chanel</td>
<td>1836</td>
<td>Gina Tricot</td>
<td>1053</td>
<td>Nine West</td>
<td>749</td>
</tr>
<tr>
<td>Levi’s(r)</td>
<td>3597</td>
<td>Gap</td>
<td>1644</td>
<td>Monki</td>
<td>957</td>
<td>Calvin Klein</td>
<td>742</td>
</tr>
<tr>
<td>Fashion Brand</td>
<td>Code</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>-------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban Outfitters</td>
<td>3450</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Steve Madden</td>
<td>1575</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OASAP</td>
<td>908</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pimkie</td>
<td>732</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New Look</td>
<td>3271</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Romwe</td>
<td>1564</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Target</td>
<td>902</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alexander Wang</td>
<td>732</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bershka</td>
<td>3169</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Michael Kors</td>
<td>1523</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FrontRowShop</td>
<td>901</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saint Laurent</td>
<td>671</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**B. Survey**

We are interested in assessing the aesthetic quality of human photos on three factors: clothing compatibility, facial attractiveness and body attractiveness. These are subjective measures, but you can guide your judgments according to the following instruction and examples.

**Instruction.**

1. **Clothing Compatibility:** how well the clothing items match together.

   **Example.**
   
   The 1st item is compatible with the 2nd item but not compatible with the 3rd item.

2. **Facial (Body) attractiveness** is the degree to which a person's facial (body) features are considered aesthetically pleasing or beautiful. Below are some general aspects you can refer to.
Examples.

Below we offer some examples of faces and bodies with different attractiveness levels (1 to 10). Please look at these examples to set your expectations.
Questions.

Please answer the following questions regarding the given picture. For the first two questions, please disregard clothing styles and focus only on face (body) attractiveness:

[Compatibility] Do you think the clothing items in the photo are compatible (match well)?
Yes/No

[Body attractiveness] How attractive do you think the person’s body is? (1: least attractive, 10: most attractive)
(Scale 1~10)
[Facial attractiveness] How attractive do you think the person’s face is? (1: least attractive, 10: most attractive)
(Scale 1~10)

[Picture aesthetics] How visually pleasing do you think the whole picture is? (1: least visually pleasing, 10: most visually pleasing)
(Scale 1~10)

[Clothing Fashion] How fashionable is the clothing style? (1: worst fashion, 10: best fashion)
(Scale 1 ~10)

[Clothing Price Appearance] How expensive do you think the clothings are? (1: very cheap, 10: very expensive)
(Scale 1 ~10)

[Subject Gender] Is the person in the picture a female or male?
Male/Female

[Subject Glasses] Is the person wearing glasses in this picture?
Yes/No

What is your age?
A: <20
B: 20~29
C: 30~39
D: 40~49
E: 50~60
F: >60

What is your gender?
Female/Male
What is your ethnicity?
A) White
B) Black or African American
C) East Asian (e.g., Chinese, Korean, Japanese)
D) Other Asian
E) Hispanic or Latino
F) Other

C. Solve for optimal style choices

For a given targeted $\overline{\text{Like}}_t$, which affects the future utility, each individual would choose styles that maximize their current utility. That is, user $i$ solves the following optimization problem (ignoring the subscript $i$):

$$\max_{(f_1, f_2)} \left[ \beta f_1^{\rho_2} + (1 - \beta) f_2^{\rho_2} \right]^{\frac{1}{\rho_2}}$$

s.t. $b_1 f_1 + b_2 f_2 = L$

where $L$ is a constant given by

$$L = \overline{\text{Like}}_t - \tilde{b} \cdot (x_1, x_2, x_3, x_1x_2, x_1x_3, x_2x_3, \text{Like}_{t-1}, \text{face}, \text{body}, \text{age}, \text{gender})$$

and $\tilde{b}$ is the coefficients estimated from the state transition regression $\tilde{g} (\cdot)$.

For the Lagrangian

$$H = [\beta f_1^{\rho_2} + (1 - \beta) f_2^{\rho_2}]^{\frac{1}{\rho_2}} + \lambda [L - (b_1 f_1 + b_2 f_2)]$$

The first order conditions are

$$L_{f_1} = \frac{1}{\rho_2} [\beta f_1^{\rho_2} + (1 - \beta) f_2^{\rho_2}]^{\frac{1}{\rho_2}} \beta \rho_2 f_1^{\rho_2-1} - \lambda b_1 = 0$$

$$L_{f_2} = \frac{1}{\rho_2} [\beta f_1^{\rho_2} + (1 - \beta) f_2^{\rho_2}]^{\frac{1}{\rho_2}} (1 - \beta) \rho_2 f_2^{\rho_2-1} - \lambda b_2 = 0$$

Solving the above system of equations, we have
\[
\frac{b_1}{b_2} = \frac{\beta f_1 \rho_2^{-1}}{(1 - \beta) f_2 \rho_2^{-1}}
\Rightarrow
f_2 = f_1 \left[ \frac{b_2 \beta}{b_1 (1 - \beta)} \right]^{\frac{1}{\rho_2 - 1}}
\]

Plug in \(L_2 = 0\), we have

\[
f_1^* = \frac{L}{b_1 + b_2 A}, \quad f_2^* = \frac{L \cdot A}{b_1 + b_2 A}
\]

where \(A = \left[ \frac{b_2 \beta}{b_1 (1 - \beta)} \right]^{\frac{1}{\rho_2 - 1}}\).

**D. Estimation Algorithm.**

Let \(I\) denote the total number of bloggers, \(N\) is the number of previous iterations used for calculating the expected value for the current iteration.

1. At iteration \(r\), the state-depend value \(V\) and heterogenous parameters \(\Theta_i\) in the past \(N\) iterations are \(H^r = \{\Theta_i^m, \tilde{V}^m(S_i^m; \Theta_i^m)\}_{m=1}^{r-1}\), where \(\theta_i \in \Theta_i = \{\alpha_{i1}, \alpha_{i2}, \rho_i, \eta_i, t_i, y_i, \delta_i\}\).

2. Draw \(\mu^r_\theta\), the population mean of \(\theta_i\), from the posterior distribution based on \(\sigma^r_\theta^{-1}\) and parameters estimated in the last iteration \(\{\theta^r_{i-1}\}_{i=1}^{L}, i.e., \mu^r_\theta \sim N \left( \frac{\sum_{i=1}^{L} \theta^r_{i-1}}{L}, \sigma^r_\theta^{-1} \right)\).

3. Draw \(\sigma^r_\theta\), the population variance of \(\theta_i\), from the posterior distribution based on the updated \(\mu^r_\theta\) and \(\{\theta^r_{i-1}\}_{i=1}^{L}\), i.e., \(\sigma^r_\theta \sim IG \left( \frac{L}{2}, \frac{\sum_{i=1}^{L} (\theta^r_{i-1} - \mu^r_\theta)^2}{2} \right)\).

4. Draw new parameters \(\theta^r_i\) for each individual \(i = 1, ..., I\) from the posterior distribution \(f_i(\theta^r_i | \mu^r_\theta, \sigma^r_\theta, S_i^0, B_{r_i}^0, P_{i}^0) \propto \pi(\theta_i | \mu^r_\theta, \sigma^r_\theta)L \{S_i^0, B_{r_i}^0, P_{i}^0 | \theta^r_i\}\), where \(S_i^0, B_{r_i}^0, P_{i}^0\) are the observed choices of style, brand, and post.

We use Metropolis-Hastings algorithm to draw from the above posterior distribution.
(1) Draw candidate parameters $\theta_i^{tr}$ from a proposal density $q(\theta_i^{tr-1}, \theta_i^{tr})$, essentially adding some perturbation to $\theta_i^{tr-1}$, for example, $\theta_i^{tr} \sim N(\theta_i^{tr-1}, \epsilon^2)$.

(2) Given $\theta_i^{tr}$, compute the likelihood, i.e., $L(St_i^o, Br_i^o, P_i^o | \theta_i^{tr})$. Computation of the likelihood with the observed continuous choice ($St_i^o$) is traditionally done in two ways: first, discretizing the continuous choice space into countable discrete choices; second, numerical approximation using kernel smoothing (e.g., Yao et al., 2012; Liu et al., 2018). The downside of the first way is the loss of the continuous nature of the corresponding choice. The second approach reserves continuity but requires thousands of draws of random errors and solving for optimal choices, which could be computationally very costly, especially for the case without closed-form solutions.

Our approach borrows the spirits of both ways. On the one hand, we reserve the continuous nature of the style choices $St_i^o$, that is, we allow the choices to take any positive real numbers. On the other hand, to alleviate the computational burden, we also use some numerical approximation or smoothing, based on the ‘discretizing’ the number of likes which naturally take discrete values (i.e., integers).

Specifically, for a given target likes $\hat{\text{Like}}$, there exists a unique choice $St_i^*$ (please see Appendix C for the optimal solution). The $\hat{\text{Like}}$ can take any integer from 0 to the maximum of likes achieved across all the posted fashion looks $K = 1747$. The probability for the observed style choices $St_i^o$ that lead to $\hat{\text{Like}} = \hat{g}(St_i^o, Br_i^o, S_{it})$ is therefore

$$Pr\{St_i^o, Br_i^o, P_i^o | S_{it}, \theta_i^{tr}\} = \exp \left\{v^r(\hat{\text{Like}}, Br_i^o, P_i^o | S_{it}, \theta_i^{tr})\right\}$$

$$= \frac{\exp \left\{v^r(\hat{\text{Liker}}, Br_i^o, P_i^o | S_{it}, \theta_i^{tr})\right\}}{\sum_{k=0}^{K} \sum_{n=1}^{35} \exp \left\{v^r(k, Br_i, P_i^o = 1 | S_{it}, \theta_i^{tr})\right\}}$$

where $v^r(\hat{\text{Liker}}, Br_i^o, P_i^o | S_{it}) = v^r(St_i^o, Br_i^o, P_i^o | S_{it})$ if $\hat{\text{Like}} = \hat{g}(St_i^o, Br_i^o, S_{it})$ is an integer. Otherwise, $v^r(\hat{\text{Liker}}, Br_i^o, P_i^o | S_{it}) = v^r(\text{Like}_nbr, Br_i^o, P_i^o | S_{it}) = v^r(St_i^*, Br_i^o, P_i^o | S_{it})$, where $\text{Like}_nbr$ is the nearest integer neighbor of $\hat{\text{Like}}$, and $St_i^*$ is the optimal choices to achieve the target $\text{Like}_nbr$.

The choice specific value function $v^r(St_i^o, Br_i^o, P_i^o | S_{it})$ is the per-period utility plus the expected future value $EV^r(S_i)$, calculated with a weighted average of the past state-specific
value \{(V^m(S^m_l; \Theta^m_i))_{i=1}^r\}_{m=r-N}. The cumulative number of likes, denoted by s, as the stochastically evolving state, follow the state transition probability \(T(St_i, Br_i, P_i|s, \sigma)\). We have the estimated standard deviation for the random error \(\zeta_i \sim N(0, \sigma^2)\), resulting from the transition regression. With a step size of 1, the probability of getting \(y\) likes rather than \(L\) is therefore 
\(1 \times \phi(0, \sigma)(y - L)\), where \(\phi(0, \sigma)(\cdot)\) is the density function for normal distribution \(N(0, \sigma^2)\). So, we have

\[
EV^r(St_i, Br_i, P_i, S_i) = \sum_{m=r-N}^{r-1} \varphi^m(S^m_l; \Theta^m_i) \cdot \frac{k_h0(\theta_i^r - \theta_i^m)k_h5(s_i^m - s_i)}{\sum_{j=r-N}^{r-1} k_h0(\theta_j^r - \theta_i^m)k_h5(s_j^m - s_i)}.
\]

Then we can calculate the likelihood \(L(St_i^0, Br_i^0, P_i^0|\theta_i^r)\).

(3) Repeat the above to obtain the likelihood for the old parameter from the last iteration \(L(St_i^o, Br_i^o, P_i^o|\theta_i^{r-1})\).

(4) Having obtained the likelihood, we can determine whether to accept the candidate parameters \(\theta_i^r\), with acceptance probability \(\lambda\) given by

\[
\lambda = \min \left\{ \frac{\pi(\theta_i^r | \mu_i^0, \sigma_i^0) L(St_i^0, Br_i^0, P_i^0 | \theta_i^r) q(\theta_i^r, \theta_i^{r-1})}{\pi(\theta_i^{r-1} | \mu_i^0, \sigma_i^0) L(St_i^0, Br_i^0, P_i^0 | \theta_i^{r-1}) q(\theta_i^{r-1}, \theta_i^r)} , 1 \right\}.
\]

Repeat (1) ~ (4) for all individuals \(i = 1, ..., I\).

5. Given the accepted parameters resulting from step 4, \(\theta_i^r\), we can update the value function for each individual, \(\{V^r(S^r_l; \Theta^r_i)\}_{i=1}^I\).

Given the Type-1 extreme value distribution for the brand-and-style specific random error, the value takes the following form

\[
V^r(S^r_l; \Theta^r_i) = 0.577 + \log \left[ \exp \{v^r(P_i^0 = 0|S_l, \theta_i^r)\} + \sum_{k=0}^{K} \sum_{n=1}^{35} \exp \{v^r(k, Br_t, P_i^0 = 1|S_l, \theta_i^r)\} \right].
\]

6. Proceed to the next iteration \(r + 1\) and repeat the above steps until convergence.

\[\text{An ordinal logistic regression was also tested for robustness check.}\]