Paying the Pink Tax on a Blue Dress - Exploring Gender-based Price-Premiums in Fashion Recommendations

Alexander Brand[0000-0003-3171-9619] and Tom Gross[0000-0001-8353-7388]

Human-Computer Interaction Group, University of Bamberg, Kapuzinerstr. 16
96047 Bamberg, Germany
hci@uni-bamberg.de

Abstract. In the field of Human-Computer Interaction there is considerable awareness on diversity and inclusion. At the same time topics such as gender and race have become more prominent recently. One aspect that has received little attention, however, is the possible reproduction of real-world socio-demographic inequality structures through recommendation systems in fashion. To investigate gender-specific differences in recommender systems, we utilise data from Amazon and use quantile regressions to calculate what price differences exist for the recommended products concerning the primary product. Our results show a bias in recommended pricing premiums about addressed gender. While a higher price in comparison to the viewed product is charged for all genders, product recommendations for women generally show a higher premium than those for men (about 5% more at the median, ceteris paribus). This can be influenced by the starting price and the popularity of the product, i.e. the sales ranking.
Keywords: GenderIT, Gender Price Gap, Recommendation Bias.

1 Motivation

Product recommendations are a relevant part of the shopping experience when buying clothes virtually [1]. The possibility to access similarly presented products in addition to the directly found items and to compare them simply and playfully offers advantages for users as well as for suppliers [2]. Sellers can present a wider range of products and potential buyers get better insights. In this respect, it is central for a good user experience that recommendations represent balanced alternatives. However, the question arises whether the underlying recommendation systems work in a non-discriminatory way as desired. From an HCI perspective this is especially important, because such factors undermine the idea of building systems that are equally accessible and usable for all. Here, aspects like gender [3, 4] and race [5] are becoming more and more part of the discourse and shape the thinking about what is socially expected when designing software. Accordingly, research on discriminatory patterns in established shops is necessary to learn about potential sources of bias and where to start when trying to minimise it, when engineering new software. One factor that has become prominent in recent years is the so-called "pink tax" [6-9], i.e. the fact that products for women as consumers have higher prices than their male equivalents.

This factor is observed especially with popular products [10]. For example, a study by the consumer advice centre in Hamburg, Germany, showed that they found such patterns in online stores like Douglas and dm (drug store chains) [11]. This raises the question of whether recommendations on such sites also follow or even reinforce such pricing patterns or counteract this trend. In this article, we use a large product data set from Amazon [12, 13] to investigate to what extent such tendencies can be identified for the clothing industry. Our evaluation uses quantile regression models [14] to take into account the wide dispersion of the price premium and to obtain adequate estimates for the parameters. In this context we address the following research questions:

Q1) How do price differences of recommendations to clothes differ according to the gender addressed?

Q2) How does the interplay between addressed gender and popularity vary in terms of price?

Q3) How does the interplay between addressed gender and original price vary in terms of price?

2 Related Work

Concerning the effects of gender-oriented pricing, Stevens et al. [8] show that despite the awareness of unequal treatment about prices, women are willing to buy them if the product signals sufficient affection. A comprehensive report on the extent of differences is provided by An der Heiden und Wersig [11] who find considerable differences in services such as haircutting. Here, only 11% of all providers would offer men and women the same prices for short haircuts. Also for clothes, specific offers like professional cleaning 68% would charge different prices. The authors provide similar results for the German market as previous studies on price differences for the USA and UK [15, 16]. Concerning possible bias in recommender systems, Ekstrand et al [17] show for the online book market that recommendations can have a gender bias, but differences depending on the underlying algorithm have to be taken into account. In principle, however, recommender systems could also help to counteract filter bubbles and not support them, as the work of Nguyen et al shows [18]. Such an effect is particularly relevant when personalised recommendations are considered, as these recommendations are of particular importance, also in comparison to recommendations closer to the people, such as "other customers also bought", as considered in this study. In relation to more general approaches, the branch of the investigation of fairness in recommender systems [19], which methodically deals with the reduction of bias has emerged. Depending on the area, bias can refer to socio-demographic characteristics [20] as well as undesirable patterns due to strong differences in popularity [21]. Possible approaches are punishing the corresponding model parameters if the bias is too strong [19]. While these approaches include a variety of areas of application the authors are not aware of any in-depth statistical analysis of gender bias for the field of fashion, which is especially interesting because pricing itself is analysed quite often [22-25].

3 Theoretical Background

Pink Tax describes a circumstance, which assumes that a similar or the same product is offered with different prices for men and women. Often products for women are priced higher than those for men [10]. In addition to a price component, this is also reflected in the product range—for instance, when the choice of products for women is smaller [7]. A combination of signalling [10] and role theory [26, 27] can be used as a justification for the perceived use. Looking at signalling, it can be explained that the products offered are interpreted by the female buyers in such a way that the annotations made to the products give the impression of a gender connotation of the product [10]. Consequently, the purchase of pink products or objects "for Her", as an example, illusorily illustrates the affiliation to the group of "real" women [13]. Consequently, the use of a gender-addressed object represents an act of legitimation of personal sexuality [10]. Duesterhaus et. al. conclude that correspondingly larger price differences should be found for publicly visible aspects such as hairstyles and feminine clothing [10]. Furthermore, when considering gender pricing, the specific role and public display of women [28] in the social and working environment must be taken into account. For example, the fear of potential disadvantages that may arise from the inadequate use of gender-specific, reputable clothing may encourage purchasing decisions for correspondingly gender-addressed products [10]. This can be applied to the consideration of additional recommendations in a way that the alternatives to a product define a range of acceptable alternatives. These are legitimised by the fact that the alternatives showed are either similar in nature, compatible, or serve the same purpose. Since the alternatives define a price range, they indicate which range of prices is acceptable for a given item. According to the expectations regarding the individual prices, products, which address women, should also show a right-shifted distribution of the prices of the recommended products. If these are popular products, there should also be a larger price difference, because it would be expected that these products would be used to meet social expectations and would therefore be bought at higher premiums. Consequently, a larger price premium can be expected for products addressed to women. Additionally, due to gender pricing for more everyday products, a larger difference should be found for more frequently demanded goods, i.e. in this case more popular items. In this respect, the following hypotheses are to be tested:

H1) Products for women have more expensive recommendations than products for men.

H2) The price difference is influenced by the popularity of the product.

H3) The price difference is influenced by the price of the product.

4 Methods

For the evaluation, Amazon product data was used because the recommendations are presented explicitly and separately based on the current product under consideration. For the analyses, the recommender system data of UC San Diego [12, 13] with 1,503,384 products was used, which include data such as a description of the item, price, product categories, brand, sales rank, and alternative clothing items recommended for the product. We calculated for all products the ratio from the average of the prices of the recommended products and the displayed products and named the result “Price Premium”. In this respect, the target value is a variable that has the value range $[0, +\infty]$. Values > 1 indicate a more expensive recommendation, while values < 1 mean that the recommended products are cheaper than the selected product. Besides, information on the products was extracted from the stored product categories. Since clothing for young adults and children can also have other rules for consumption [29], only the categories "Women" and "Men" were used as indicators for the gender assignment of the products. Due to non-normality ($p < 0.001$, Shapiro-Wilk-Tests), non-parametric methods were chosen. To test the hypotheses, quantile regression models were used, which on the one hand allow to relax the underlying distribution assumptions and on the other hand are robust against outliers and thus are often applied in pricing research [14, 30-32]. For the calculation, interaction effects between the addressed gender and the sales rank were taken into account. Furthermore, the price of the original product and interactions of the original price with the addressed gender were considered in the modelling to test for general pricing effects. In this respect the unrestricted model can be written as:

$$\begin{aligned} Q_\tau (Premium_i) = {} & \beta_0(\tau) + \beta_1(\tau)\, Gender + \beta_2(\tau)\, Rank + \\ & \beta_3(\tau)\, Gender * Rank + \beta_4(\tau)\, Price + \beta_5(\tau)\, Gender * Price + \varepsilon \end{aligned} \tag{1}$$

Where τ is considered for the values 0.1 - 0.9 and $\beta_0(\tau), \beta_1(\tau), \ldots, \beta_5(\tau)$ are calculated accordingly as the result of a minimization problem for the estimated value of τ.

5 Results

Regarding the used variables we observe a high dispersion in Price, Price Premiums, and Ranking with a high right skewness. The mean price is around 34$ and the mean price premium is around 1.28 which indicates a surplus of about 28% for recommendations. Regarding the addressed gender we observe 244908 products for women and 140117 products for men. The descriptive analysis of the variables addressed gender, sales rank, price and average recommended price show bivariate relationships of varying strength and direction. If we look at possible differences according to gender, significant differences between the addressed genders can be seen for both the price and the price difference between product and recommendation. A Mann-Whitney U test shows that the median of products for women ($Mn = 22$) surpass those for men ($Mn = 21.6$) on a significant level ($W = 3649143244$, $p = 0.0004$). The same is observable for the distribution of the price premiums. Accordingly, a Mann-Whitney U test shows that female products ($Mn = 1$) surpass products, which address men ($Mn = 0.988$) on a significant level ($W = 2622305840$, $p < 2.2e\text{-}16$). The usage of quantile regression lines reveals a negative relationship in the graphical representation of ranks against premiums. It is important to note here that this effect can only be differentiated for ranks > 100 and that the sample contains too few products addressed to women. Otherwise the pure plotting of the observations does not show a clear direction when ignoring the addressed gender.

Testwise, we observe the same effect as in Figure 1 (right side). Our calculations using spearman rank correlation tests indicate that there is a significant relationship between the premium and the sales rank ($r_s = .13$, $p < .001$).

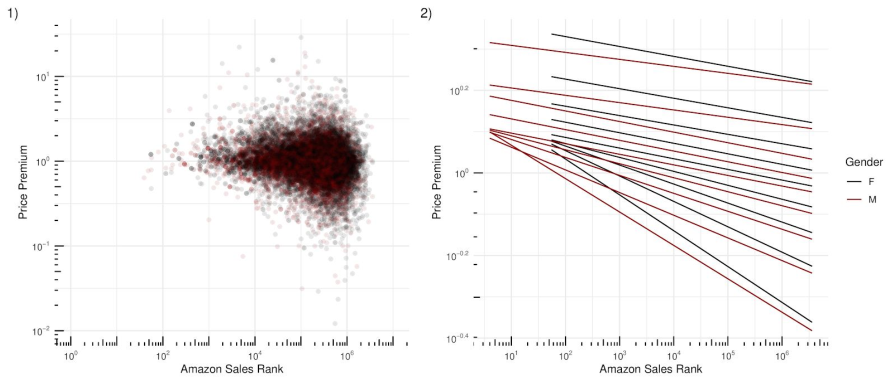

Fig. 1. Relationship between (Sales) Rank and Price Premiums. Left side displays the distribution of observations on a log-log scale to account for high dispersion. Right side shows the results of gender-facetted quantile regression smoothening. The lines symbolise the 10 % quantiles within a gender (F = Female, M = Male), which are displayed in different colours.

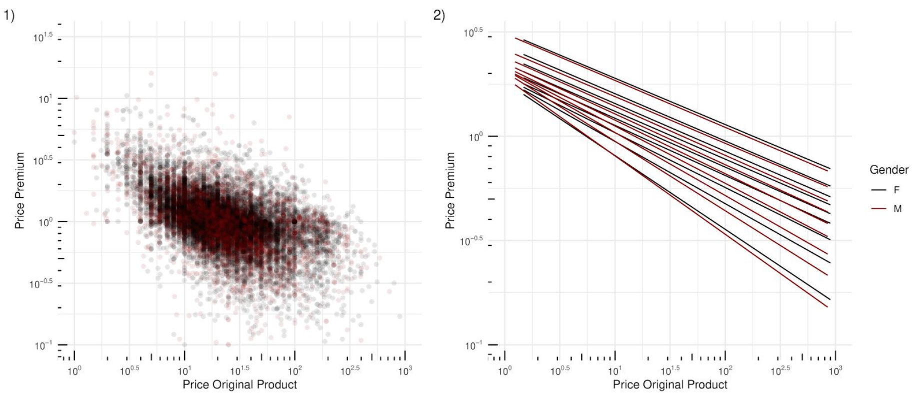

Fig. 2. Relationship between Price and Price Premiums. Left side displays the distribution of observations on a log-log scale to account for high dispersion. Right side shows the results of gender-facetted quantile regression smoothening (F = Female, M = Male) like figure 1.

Looking at the regression model the available number of observations for the test variables differ highly, which leads to a usable sample of 150919 observations for the model. For $\tau = 0.5$ (median), a significant effect on the parameter "For Men" ($\beta = -0.014$, $p<0.001$) can be observed. Considering the other variables, continues to increase and finally has a value of $\beta = -0.053$ ($p<0.001$).

This means higher prices for female targeting products ceteris paribus. A positive significant effect ($p < 0.001$) can be shown for the interaction with the starting price, which has an effect of β = -0.0004, which shows a dependency of the gender effect in relation to the original product price. Looking at the variable Sales Rank (1000s) a negative significant relationship (β = -0.001, p<0.001) can be observed. This means unpopular products tend to have lower gendered differences in price premiums ceteris paribus. A similar picture is shown when the calculation is carried out for all values of τ. Here, for all considered values of the parameter "For Men", an effect significantly different from 0 is shown concerning the difference between the product and recommended products. A similar effect can be observed for price and rank. A more differentiated effect can be determined for the interaction terms. Here significant positive effect can be seen for larger values of τ. This can be interpreted to mean that for very high price differences, other variables (e.g. discounts) may be more important. In order to check the influence of brands, the average brand prices and the quantiles of the average prices of the brands were checked. Here, no relevant differences were found about the parameters. In summary, the results indicate a higher price premium for female targeting products, which depends on the original price and the popularity.

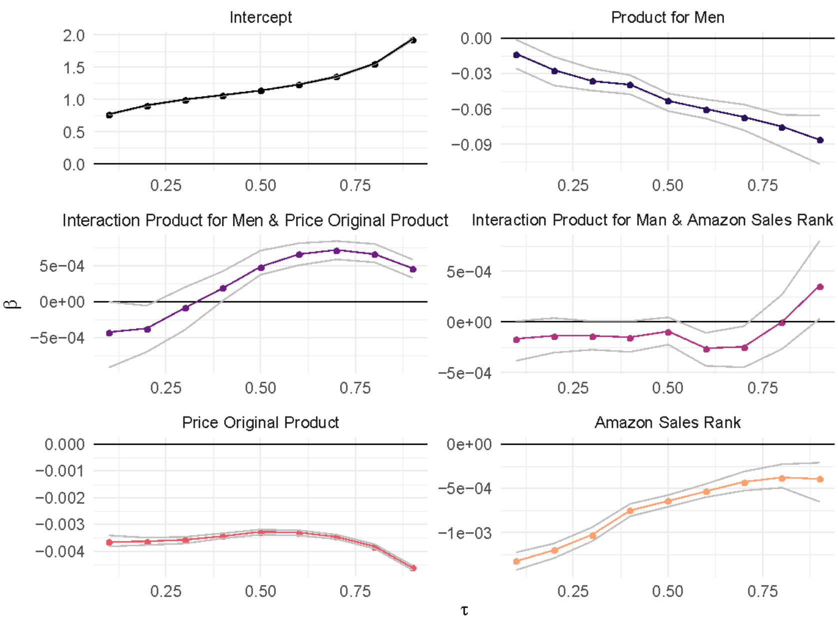

Fig. 3. Parameter estimations of β for the unrestricted model (see methods section) for given τ. The area around the estimates shows the 95% confidence band.

6 Discussion

Our results show a significant difference in the prices of recommendations depending on whether the product is addressed to men or women. The effect is slightly determined by the starting price and the popularity of the product. However, there are large differences in the quantiles of the distribution of the premium. While smaller and medium price differences according to gender are rather little influenced by the other variables considered, these factors show an influence on the gender difference, especially for higher-priced and less popular products. In this respect, the addressed gender seems to make a difference, but this is not suitable to explain extreme differences such as > 78% (90%-quantile) premium alone. Following H2, interactions have to be considered here. In this respect, the argument of a broadly observable pricing strategy differentiated according to gender show that even in the case of a large provider, corresponding discriminatory structures can still be found. Another effect is that the coefficients are not much lower for popular products, which shows that such effects are not only relevant for fringe group interests, but also meet the casual shopper. One aspect which is put forward here in defence is that products for women are not directly comparable to products for men, as there might be differences in quality or products might not be comparable due to other characteristics. In this case, the algorithm used already constitutes a space of products that can be considered as comparable. Additionally, for clothing that is mass-produced there should not normally be any major differences in production costs [11] and as checked before, the effect is not affected much by the chosen brand.

7 Limitations

We showed that fashion recommendation systems have the potential to reproduce sexist pricing strategies. Further research is needed with respect to recommendation algorithms [33, 34, 35]. Especially the consideration of the different brands suggests very different gender orientation bases, even if recommendations also include clothes outside the considered label [36]. The data set used here offers little comparable meta information about the products and buying behaviour. Further research should explore other relevant background factors in greater depth and compare between different suppliers. Besides, the focus was only on items of clothing that were explicitly identified by Amazon as for men or women. Here, image recognition and natural language processing methods can help to classify the target groups more precisely to further increase the number of cases on the one hand and to reduce potential selective failures due to designation errors on the other.

References

1. Kotouza, M.T., Tsarouchis, S.F., Kyprianidis, A.-C., Chrysopoulos, A.C., Mitkas, P.A.: Towards Fashion Recommendation: An AI System for Clothing Data Retrieval and Analysis. In: IFIP International Conference on Artificial Intelligence Applications and Innovations, pp. 433-444. Springer, (2020)

2. Senecal, S., Nantel, J.: The influence of online product recommendations on consumers' online choices. Journal of Retailing 80, 159-169 %@ 0022-4359 (2004)
3. Bardzell, S.: Feminist HCI: taking stock and outlining an agenda for design. pp. 1301-1310. (2010)
4. Stumpf, S., Peters, A., Bardzell, S., Burnett, M., Busse, D., Cauchard, J., Churchill, E.: Gender-Inclusive HCI Research and Design: A Conceptual Review. Foundations and Trends® in Human–Computer Interaction 13, 1--69 (2020)
5. Schlesinger, A., Edwards, W.K., Grinter, R.E.: Intersectional HCI: Engaging identity through gender, race, and class. pp. 5412-5427. (2017)
6. Jacobsen, K.A.: Rolling Back the "Pink Tax": Dim Prospects for Eliminating Gender-Based Price Discrimination in the Sale of Consumer Goods and Services. California Western Law Review 54, 2 (2018)
7. Lafferty, M.: The Pink Tax: The Persistence of Gender Price Disparity. 17 (2019)
8. Stevens, J.L., Shanahan, K.J.: Structured Abstract: Anger, Willingness, or Clueless? Understanding Why Women Pay a Pink Tax on the Products They Consume. Creating Marketing Magic and Innovative Future Marketing Trends, pp. 571--575. Springer International Publishing (2017)
9. Yazıcıoğlu, A.E.: Pink Tax and the Law: Discriminating Against Women Consumers. Routledge (2018)
10. Duesterhaus, M., Grauerholz, L., Weichsel, R., Guittar, N.A.: The cost of doing femininity: Gendered disparities in pricing of personal care products and services. Gender Issues 28, 175--191 (2011)
11. An der Heiden, I., Wersig, M.: Preisdifferenzierung nach Geschlecht in Deutschland. (2017)
12. McAuley, J., Targett, C., Shi, Q., Van Den Hengel, A.: Image-based recommendations on styles and substitutes. pp. 43--52 (2015)
13. He, R., McAuley, J.: Ups and downs: Modeling the visual evolution of fashion trends with one-class collaborative filtering (2016)
14. Koenker, R., Hallock, K.F.: Quantile regression. Journal of economic perspectives 15, 143--156 (2001)
15. Liston-Heyes, C., Neokleous, E.: Gender-based pricing in the hairdressing industry. Journal of Consumer Policy 23, 107-126 %@ 0168-7034 (2000)
16. Whittelsey, F.C.: Why women pay more: How to avoid marketplace perils. (1993)
17. Ekstrand, M.D., Tian, M., Kazi, M.R.I., Mehrpouyan, H., Kluver, D.: Exploring author gender in book rating and recommendation. In: Proceedings of the 12th ACM Conference on Recommender Systems, pp. 242--250. (2018)
18. Nguyen, T.T., Hui, P.-M., Harper, F.M., Terveen, L., Konstan, J.A.: Exploring the filter bubble: the effect of using recommender systems on content diversity (2014)
19. Islam, R., Keya, K.N., Pan, S., Foulds, J.: Mitigating demographic biases in social media-based recommender systems. KDD (Social Impact Track) (2019)
20. Kamishima, T., Akaho, S.: Considerations on recommendation independence for a find-good-items task. (2017)
21. Abdollahpouri, H.: Popularity Bias in Ranking and Recommendation (2019)
22. Archak, N., Ghose, A., Ipeirotis, P.G.: Deriving the Pricing Power of Product Features by Mining Consumer Reviews. Management Science 57, 1485-1509 (2011)
23. Lee, K., Lee, Y.S., Nam, Y.: A novel approach of making better recommendations by revealing hidden desires and information curation for users of internet of things. Multimedia Tools and Applications 78, 3183–3201 (2019)
24. Schoinas, I.: Product Recommendation System. (2019)
25. Zheng, Y., Gao, C., He, X., Li, Y., Jin, D.: Price-aware Recommendation with Graph Convolutional Networks. pp. 133–144. IEEE, (2020)
26. Goffman, E.: Gender display. Gender advertisements, pp. 1-9. Springer (1976)

27. Beckmann, C., Gross, T.: Social Computing—Bridging the Gap between the Social and the Technical. In: International Conference on Social Computing and Social Media, pp. 25-36. Springer, (2014)
28. West, C., Zimmerman, D.H.: Doing gender. Gender & society 1, 125-151 %@ 0891-2432 (1987)
29. Darian, J.C.: Parent-child decision making in children's clothing stores. International Journal of Retail & Distribution Management %@ 0959-0552 (1998)
30. Chen, C.W.S., Li, M., Nguyen, N.T.H., Sriboonchitta, S.: On asymmetric market model with heteroskedasticity and quantile regression. Computational Economics 49, 155--174 (2017)
31. Hung, W.-T., Shang, J.-K., Wang, F.-C.: Pricing determinants in the hotel industry: Quantile regression analysis. International Journal of Hospitality Management 29, 378--384 (2010)
32. Mak, S., Choy, L., Ho, W.: Quantile regression estimates of Hong Kong real estate prices. Urban Studies 47, 2461--2472 (2010)
33. Gross, T. Group Recommender Systems in Tourism: From Predictions to Decisions. In Proceedings of the 2nd Workshop on Recommenders in Tourism - RecTour 2017 - Co-Located with the 11th ACM Conference on Recommender Systems - RecSys 2017 (27. Aug., Como, Italien). Neidhardt, J., Fesenmaier, D., Kuflik, T. and Woerndl, W., eds. CEUR Workshop Proceedings, Vol. 1906, Aachen, Deutschland, 2017. pp. 40-42.
34. Beckmann, C. and Gross, T. Towards Increased Utility of and Satisfaction with Group Recommender Systems. In Extended Proceedings of the 21st Conference on User Modelling, Adaptation, and Personalisation - UMAP 2013 (June 10-14, Rome, Italy). Berkovsky, S., Herder, E., Lops, P. and Santos, O.C., eds. CEUR Workshop Proceedings, Vol. 997, Aachen, Germany, 2013. pp. 3-5.
35. Beckmann, C. and Gross, T. AGReMo: Providing Ad-Hoc Groups with On-Demand Recommendations on Mobile Devices. In Proceedings of the European Conference on Cognitive Ergonomics - ECCE 2011 (Aug. 24-26, Rostock, Germany). Dittmar, A. and Forbrig, P., eds. ACM, N.Y., 2011. pp. 179-183.
36. Lieven, T., Grohmann, B., Herrmann, A., Landwehr, J.R., Van Tilburg, M.: The effect of brand gender on brand equity. Psychology & Marketing 31, 371-385 %@ 0742-6046 (2014)