Investigating the Pink Tax: Evidence against a Systematic Price Premium for Women in CPG*

Preliminary and Incomplete

Natasha Bhatia† Sarah Moshary‡ Anna Tuchman§

July 7, 2021

Abstract

The *pink tax* refers to an alleged empirical regularity: that products targeted toward women are more expensive than similar products targeted toward men. This paper leverages a national dataset of grocery, convenience, and mass merchandiser sales, in combination with novel sources on product gender targeting, to provide systematic evidence on price disparities for personal care products targeted at different genders. The results are mixed: while women's deodorant products are indeed more expensive, on the order of 6%, women's disposable razors are about 8% less expensive. We then investigate potential drivers of differential pricing with a focus on deodorants. Analysis of wholesale prices indicates that differences in costs cannot fully explain the differences in deodorant retail prices that we document. Rather, our findings suggest that demand for women's deodorant products is relatively inelastic. We also find a higher share of category TV advertising features women's products. Finally, we consider the potential welfare effects of price parity regulations.

1 Introduction

The "pink tax" refers to an alleged empirical regularity that goods marketed toward women have higher prices than their counterparts marketed towards men. Gender-based

*We would like to thank Sara Drango, Malika Korganbekova, and Jason Zhao for excellent research assistance. Researcher(s) own analyses calculated (or derived) based in part on (i) retail measurement/consumer data from Nielsen Consumer LLC ("NielsenIQ"); (ii) media data from The Nielsen Company (US), LLC ("Nielsen"); and (iii) marketing databases provided through the respective NielsenIQ and the Nielsen Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the Nielsen data are those of the researcher(s) and do not reflect the views of NielsenIQ or Nielsen. Neither NielsenIQ nor Nielsen is responsible for, had any role in, or was involved in analyzing and preparing the results reported herein. Researcher(s)' own analyses calculated (or derived) based in part on data from The Numerator Company (US), LLC and marketing databases provided through the Numerator Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business." "The conclusions drawn from the Numerator data are those of the researcher(s) and do not reflect the views of Numerator. Numerator is not responsible for and had no role in analyzing or preparing the results reported herein

†Contact: natasha.bhatia@kellogg.northwestern.edu

‡Contact: sarah.moshary@chicagobooth.edu

§Contact: anna.tuchman@kellogg.northwestern.edu

Electronic copy available at: https://ssrn.com/abstract=3882214

pricing of consumer packaged goods (CPG) is concerning because it would exacerbate well-documented gender inequality in the labor market.[1] Investigative journalists and government agencies report that price differences in CPG occur most frequently for personal care products, such as deodorant and razors, and peg price differences in this category at 13%, (e.g., Bessendorf 2015; Consumer Reports 2010; Duffin 2019). Policymakers appear keen to address these perceived inequalities through legislation. For example, in 2019-2020, the NY State Assembly and State Senate passed a bill (S2679) that bans pricing for goods on the basis of gender. The bill went into effect in September 2020. In 2019, Congresswoman Jackie Speier introduced HR 2048, the Pink Tax Repeal Act, which would implement a similar ban nation-wide.

Unfortunately, there is a dearth of systematic evidence on the prevalence and causes of the pink tax to guide legislative action. The evidence on gendered price discrimination typically features contexts where price are negotiated: gender pay gaps in the labor market (e.g., Blau and Kahn 2017), price premiums in automobile sales and repairs (e.g., Ayres (1991), Ayres and Siegelman (1995), Goldberg (1996), and Busse, Israeli, and Zettelmeyer (2017)), and more recently, disparities in real estate (e.g., Goldsmith-Pinkham and Shue 2020). In these settings, a female customer may be unaware that she is quoted a higher price than male customers for the same product and/or may be unable to arbitrage to a lower price because her gender is observable to the opposite party in the transaction. In contrast, personal care products are sold in posted price markets, which do not easily sustain this type of third degree price discrimination. Indeed, the price disparities that we document do not reflect differences in prices charged to men and women for the same exact product as in the auto market, but differences in shelf price across products that are targeted at different genders. A woman can typically observe the shelf prices of products aimed at men, and there is no rule or regulation that bars her from buying a cheaper men's product. Despite these arbitrage opportunities, evidence from government and media surveys suggests price differences persist in CPG. Unfortunately, these early studies are somewhat limited: they employ hand-collected data on prices in a small geographical area at a single point in time. As an example, Bessendorf (2015), a government study frequently referenced in state and national pink tax legislation, collected prices for 122 products at 24 NYC stores and found that women's products were more expensive in six of the seven personal care categories they studied. In contemporaneous work, Gonzalez Guittar et al. (2021) finds mixed results in their study of scraped price data from four online retailers.

A first contribution of this paper is to document gender-based price differences for a wide array of personal care products at thousands of retail outlets across the United States from 2016-2018. We find that products targeted toward women are not always more expensive.

[1] Blau and Kahn (2017) provide a recent review of the literature.

Electronic copy available at: https://ssrn.com/abstract=3882214

Of the ten categories we study, women pay higher prices in only three: deodorant, bar soap, and body wash. In the other seven categories men's unit prices are higher than women's. These results contrast with the findings of more ad hoc pink tax studies publicized in the press and highlight the need for systematic evidence on pricing practices.

Our analysis departs from earlier studies in two other ways: first, we leverage data on quantities sold to speak to differences in both purchase prices and shelf prices; and second, we broaden the scope of comparison to consider the full (unconditional) distribution of prices for men's and women's goods. To be clear, the studies cited above compare prices of similar items produced by the same brand that are marketed to men and women. While this approach is useful insofar as it can control for any differences in manufacturing costs between men's and women's products, we believe this measurement methodology is narrow in scope and may overlook material differences in pricing that exist at the category level. For example, in the deodorant category, Procter & Gamble's flagship products for men and women are sold under different brand names, Old Spice and Secret, and thus are omitted from these analyses. Thus, focusing on apples-to-apples within-brand comparisons ignores the pricing behavior of some of the most important brands in personal care. This discussion of the conditional and unconditional price gaps closely parallels the evolution of the literature on the gender wage gap. Researchers have drawn important insights from both the unconditional wage distribution and the conditional wage distribution, where the unconditional distribution sheds light on the realized outcomes for each gender, while the conditional distribution deepens our understanding of why average outcomes differ across genders.

The prima facie price patterns in the data provide evidence against a systematic price premium for women's personal care products, but even in the categories where women pay more, it is not clear whether price differences should be interpreted as a "pink tax." In part, this reflects an ambiguity in the definition of the "pink tax" in whether it encompasses differences in prices due to manufacturing costs, price elasticities, firm conduct, and/or taste-based discrimination (à la Becker 1957). From a practical perspective, understanding the determinants of price differences across products targeted at different genders is key to understanding the consequences of policies that mandate price parity. We take deodorants as a case study to investigate whether the price gap reflects differences in costs or markups. In the case of deodorant, women's products are 6% more expensive per product and 33% more per unit. If these price differences reflect differences in costs, then requiring price parity might increase consumer surplus for women at the expense of total surplus. Alternatively, if price elasticities differ across genders, then eliminating the pink tax could increase total surplus. We focus on these two drivers because it is hard to understand how taste-based discrimination might affect pricing in CPG, although we cannot rule it out entirely.

A first finding is that at the wholesale level, women's deodorant products are 4% *cheaper* per product. This fact suggests that wholesale prices are not the primary driver of the price

Electronic copy available at: https://ssrn.com/abstract=3882214

gap for deodorants. Of course, wholesale prices can reflect both manufacturing costs and downstream elasticities. To estimate elasticities, we follow Hitsch, Hortaçsu, and Lin (2021): we estimate a log-log demand function for the top five women's and men's deodorant brands. Our estimates suggest that demand for women's products is relatively inelastic.

The differences in elasticities that we document explain why a profit-maximizing firm would set higher markups for women's deodorant products than for men's. While rational from a firm perspective, work by Kahneman, Knetsch, and Thaler (1986) and Anderson and Simester (2008) suggests that many consumers regard elasticity-based pricing as unfair. For some, assessing fairness may require understanding why women's demand is more inelastic and in particular, whether elasticities reflect broader social inequalities. We therefore consider several factors that could generate the elasticity differences we document. First, we find that the average store carries more products and brands targeted at men than at women. Such assortment differences could contribute to relatively inelastic residual demand for women's products and/or relatively tougher price competition for men's products. However, this finding begs a second question of why the women's deodorant market sustains fewer brands. One possibility is that advertising creates an asymmetric barrier to entry in the deodorant market. Indeed, we find that a higher share of category TV advertising is for women's products. However, we do not find significant differences in brand loyalty across male and female consumers.

Overall, our results provide evidence against a systematic pink tax in personal care products; while we find that products targeted at women are more expensive in some categories, they are less expensive in others. However, the price differences within product categories are economically meaningful, so that a consumer could reduce their personal care expenses by approximately 7% if they were willing to purchase products targeted at a different gender.[2] What sustains these price gaps? Why don't women buy men's products and vice versa? Put another way, what explains cross-price elasticities? A literature in psychology and anthropology highlights the importance of *performing gender*, taking actions or consuming products that reinforce our identity within society (for example, Duesterhaus et al. 2011). While a full analysis is beyond the scope of this paper, we note that advertising could help to reinforce the role of personal grooming products as a vehicle for gender identity as a particular kind of horizontal differentiation. Indeed, we find that almost 80% of deodorant advertising features women's products.

2 Data

Retail Prices

We use the Nielsen Retail Scanner data from 2006 to 2018 to document price differences

[2]Based on average household spend on personal care products in the 2018 Nielsen Consumer Panel dataset.

Electronic copy available at: https://ssrn.com/abstract=3882214

between personal care products targeted at men and women. We examine six categories: body wash, deodorant, hair coloring,[3] razors, shampoo and conditioner, and shaving cream. The data records the price and quantity sold for products (UPCs) sold in 39,697 stores affiliated with 93 chains across the US. The data is recorded at the store-UPC-week level, so that prices reflect the weekly average price paid by consumers in a particular store during a particular week. The data also includes product characteristics, such as the brand name, a product description, and product size.

The data does not indicate the price of a product in weeks when it earns no sales at that store. These missing prices are not problematic for our analyses that focus on average price paid. However, we must impute prices for our analysis of shelf prices and demand elasticities. First, we assume that a product with zero sales in a particular week was not on discount, but instead was offered at its regular shelf price. Then we impute prices based on adjacent weeks when the product was sold, in an approach similar to Hitsch, Hortaçsu, and Lin (2021). Web Appendix A details the algorithm that we use to construct regular (non-discounted) shelf prices.

Gender

A key ingredient in estimating the pink tax is determining the "gender" of a product. What does it mean for a product to be gendered? A first definition is products that firms explicitly target to either men or women are gendered products. To measure this type of targeting, we extracted information on gender from the following sources:

1. **Nielsen Brand and Product Module Descriptions:** We search for gendered words in Nielsen's brand description for each UPC. These words include: "his," "hers," "men," "women," and "lady."[4] Nielsen also categorizes hair coloring and shaving cream products into separate product modules by gender. As a result, we assign a gender to all of the products in these categories.

2. **Label Insight:** We collected data on gender targeting from Label Insight. Label Insight is a market research firm that records granular product attribute information for CPG brands. The database includes pictures of products and fields indicating the products' ingredients and other attributes. In particular, the data includes a gender field.[5]

[3]We exclude temporary, costume hair coloring products. These account for 1.8% of category market share. Additionally, 92% of category market share is for hair coloring products measured in counts, and the remainder is measured in ounces. To simplify our per unit analysis, we exclude hair coloring products measured in ounces.

[4]To classify men's UPCs, we searched for the following words: men, clubman, hombre, man, man cave, homme, men's choice, men's select, monsieur, and Mr. To classify women's UPCs, we searched for the following words or abbreviations: lady, girl, ldy, women, femme, ladies, lady's, and wmn.

[5]The gender field is populated for 12% of the deodorants observed in the Label Insight data.

Electronic copy available at: https://ssrn.com/abstract=3882214

3. **Walgreens Website:** We scrape gender categorizations from `Walgreens.com`, the website of the large American drugstore chain. Walgreens conveys categorizations to its site visitors in two ways. For example, it allows users to browse deodorants & antiperspirants that are "for men" or "for women." It also allows visitors to filter on gender on its Search Results Pages. Web Appendix B displays screenshots of these classifiers. Scraping was performed in Summer 2020.

To measure explicit targeting, we construct a gender variable that combines (1)-(3) above.[6] We think of this type of targeting as akin to third degree price discrimination because products are explicitly labeled by gender. Thus, consumers might face a social stigma if they purchase products targeted at another gender. However, it is possible that consumers arbitrage price differences as consumers are not required to prove their gender to purchase a product.

Of course, firms may target a product at a particular gender, even absent an explicit (verbal) label. Firms could second degree price discriminate by designing products that hold different appeal to men and women. For example, the color, size, and scent of a product might induce differential purchasing across consumers by gender. We measure this type of targeting using two additional sources:

4. **Differential purchasing by all-male and all-female households in the Nielsen consumer panel dataset from 2006 to 2018.** We utilize the Nielsen Consumer Panel dataset on household-level purchases of personal care products to identify products whose consumer base is significantly skewed towards one gender. We focus on purchases of single-gender households, which account for over a quarter of households in the panel. Female-only households are more common – they represent 71% of single-gendered households. We aggregate the data to the UPC-household gender-level. An observation in the aggregated data records the total number of households of each gender that ever purchase a given UPC. For each UPC, we define the female (male) share as the percent of total single-gender household purchases that are made by female (male) households. Finally, we identify women's (men's) UPCs as those whose female (male) share is significantly larger than that gender's representation in the panel.[7] That is, we perform a binomial test where the null hypothesis is that the female (male) share is equal to 71% (29%). If we do not reject the null, the product is left uncategorized. This strategy identifies a different set of products than the previous strategies. Rather than labeling UPCs that are explicitly marketed towards a gender, this strategy identifies UPCs that are predominantly purchased by that gender. Empirically, these two sets turn out to be similar but not identical.[8]

[6] There are no conflicts in gender assignment from these three sources.

[7] One benefit of this approach is that it does not categorize UPCs with few household purchases.

[8] For example, the brand *Old Spice* is solely marketed towards men, but it produces deodorants in scents like Fiji

Electronic copy available at: https://ssrn.com/abstract=3882214

5. **Hand-coding Label Insight product images.** As mentioned previously, the gender field is not populated for all of the products in the Label Insight data. However, all products have images available. We hired undergraduate research assistants at the University of Chicago and Northwestern University to look at the images and hand code the gender for all of the deodorants in the Label Insight data. Products were categorized as either male, female, unisex, or unknown. Web Appendix C describes the recruiting and labeling process in detail.

Finally, we construct a single gender categorization by combining all of these sources. In the event of conflicts, we prioritize the classification from the RMS brand description and break remaining ties using majority rule or in the case of even ties, the authors' judgement.[9] In a final step, we fill in the gender for unclassified UPCs for which the corresponding brand or brand-size pair has i) at least 10 UPCs in the data and ii) at least 20% of those UPCs are labeled unanimously as a single gender.

Table 1 shows the pervasiveness of gender targeting across product modules.[10] Consistent with the focus on personal care in the media surrounding the pink tax, we find that personal care categories are highly gendered. Our methods assign a gender to 37% of the products in the Nielsen data, although there is considerable variation across categories. Across categories there is also substantial variation in the share of products targeted at men vs. women; for example, the overwhelming share of hair coloring products are targeted at women, but most shaving creams are targeted at men. While the fraction of gendered products is modest in many categories, this is largely driven by the difficulty of labeling niche products. Market shares reveal that gendered products account for 80% of volume sales across all categories.[11] Table 1 also motivates our deeper dive into deodorant. A majority of deodorant products are gendered, and almost all volume sales are for gendered products. Further, the split between men's and women's products is close to even in deodorant.

---

and Citron that are also purchased by women in the Nielsen Consumer Panel dataset. However, we do find that most panelists buy products that are marketed toward their own gender. Using one person households in the Nielsen Consumer Panel, we find that 78% of female and 81% of male panelists purchase deodorants for their own gender, where product gender is defined using all sources in this section. We draw a similar conclusion if we exclude the categorization using the Panelist data.

[9] Conflicts between sources are rare. For example, less than 0.1% of deodorant products have a conflict.

[10] Nielsen has separate product modules for men's and women's hair coloring and shaving creams. We extract gender information from these categorizations and then combine the gendered product modules together.

[11] We note that a relatively high share (32%) of disposable razors sold are not assigned a gender. This lower share may be driven by the dominance of private label products (27% market share, as shown in Web Appendix G). To protect the identity of retailers in the data, Nielsen masks the UPC of private label products, so we cannot map these products to our data sources for gender.

Electronic copy available at: https://ssrn.com/abstract=3882214

Table 1: Gender Targeting across Personal Care Categories

Nielsen Product Module	% Qty Gendered		Total	% UPCs Gendered		Count of
	All	for Women	Qty (MM)	All	for Women	UPCs
Soap - Bar	71.8%	57.5%	516	19.4%	55.9%	3,438
Soap - Liquid	45.9%	94.4%	510	14.5%	90.4%	2,816
Soap - Specialty	78.5%	63.2%	822	26.5%	59.2%	6,309
Deodorants - Personal	99.0%	49.7%	1,060	76.4%	46.4%	2,683
Hair Coloring	100.0%	88.7%	307	100.0%	95.7%	2,231
Razor Blades	86.2%	33.2%	94	52.7%	37.8%	476
Razors Disposable	67.8%	51.4%	287	37.4%	50.1%	916
Razors Non-Disposable	88.8%	45.7%	69	49.6%	36.2%	417
Creme Rinses & Conditioners	80.7%	99.5%	554	26.3%	93.4%	5,384
Shampoo						
–Aerosol/Liquid/Lotion/Powder	70.6%	74.8%	896	27.1%	67.2%	7,016
–Bars/Concentrates/& Creams	66.6%	89.3%	9	22.0%	89.3%	127
–Combinations	51.1%	99.8%	20	17.9%	86.4%	452
Shaving Cream	100.0%	24.9%	269	100.0%	21.7%	870

Notes: This table describes the share of products available at Nielsen RMS stores between 2016-2018 that we record as gendered. Gender targeting is determined based on data from five sources: gendered words in the category, brand, or product description; gender claims in Label Insight; gender labels on Walgreens.com; differential purchasing by all-male and all-female households in the HMS panel; and hand-coding by undergraduate research assistants at the University of Chicago and Northwestern University.

Electronic copy available at: https://ssrn.com/abstract=3882214

Table 2: Gender Categorization by Method

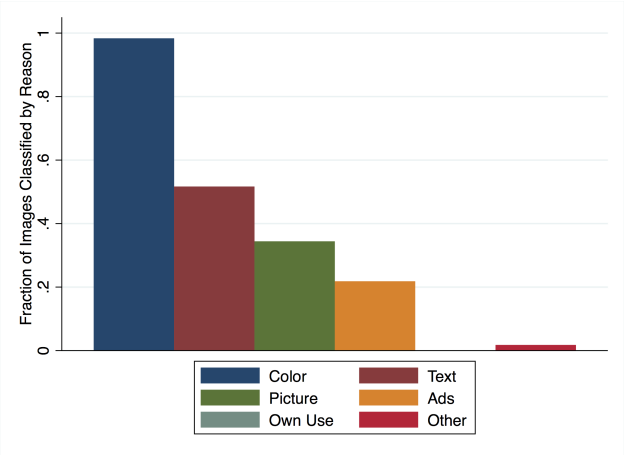
	Number of UPCs Categorized		
	Male	Female	Unisex
Third Degree			
UPC Description in RMS	232	589	0
Label Insight	137	56	3
Walgreens.com	224	230	0
Second Degree			
Single-Gender Households	1,249	1,106	0
Hand-Coded Images	573	484	0
All Sources	1,609	1,771	8
All Sources - Fill w/in Unanimous Brands	3,354	2,934	8

Notes: This table indicates the number of deodorant products sold between 2006-2018 that are categorized as gendered by each source. Gender targeting is determined based on data from five sources: gendered words in the category, brand, or product description; gender claims in Label Insight; gender labels on Walgreens.com; differential purchasing by all-male and all-female households in the HMS panel; and hand-coding by undergraduate research assistants at the University of Chicago and Northwestern University. The All Sources row represents the aggregation across these five classification sources. Only 26 products have a conflict across sources. Conflicts were broken by majority rule and in cases of even ties, by the authors' judgement.

To illustrate the contributions of our different sources, Table 2 provides a breakdown of gender categorizations by source for deodorants. The household panel data, hand-coded Label Insight pictures, and Nielsen brand descriptions are the sources that classify the most products. To get a sense for how gender targeting cues are communicated to consumers, we asked three of the research assistants to indicate which factor(s) determined their categorizations when hand-coding Label Insight images: (a) product or packaging color, (b) text on the product or packaging, (c) picture or image on the product or packaging, (d) advertisements for the product seen on television, magazines, or billboards, (e) prior knowledge of the brand or products, including own use of the product, or (f) other. Figure 1 shows the share of products for which at least one RA cited each of these factors. For over 98% of products, at least one RA reported that color influenced their understanding of gender targeting–color was by far the most dominant form of gender communication.

Electronic copy available at: https://ssrn.com/abstract=3882214

Figure 1: Factors that Communicate Gender Targeting



Notes: Based on gender categorization by three research assistants who categorized 962 deodorant products.

Wholesale Prices

We use PromoData on wholesale costs, which aggregates data from 12 major grocery wholesaler organizations in 29 markets across the United States. PromoData states that the data is collected from one major wholesaler in each market and "since price-changes and deal offers must be announced to all factors in a market at the same time, one source in each market provides a good sample on the basis of all markets reported." The dataset is primarily marketed to industry players interested in analyzing their competitors' promotion and pricing decisions. Data is reported at the UPC-market-year level. Our dataset spans 2006-2011, 30 markets, and 981 UPCs. The data includes regular wholesale prices and deal prices that reflect manufacturer incentives.[12] We use price fields that are normalized by package size and prices per package.

Advertising

Nielsen's AdIntel database provides information on advertising occurences from 2010 onwards. We merge ad occurences with viewership data to yield estimated advertising gross rating points (GRPs), the percentage of all TV-viewing households that were exposed to the ad. We aggregate the advertising data to the product level by summing GRPs across occurrences, allowing us to merge the advertising data with the RMS sales data.

[12]Deal types include advertising allowances, promotion allowances, and display allowances.

Electronic copy available at: https://ssrn.com/abstract=3882214

Demographic Data

We supplement the store-level scanner data with demographic data, which we use to explore whether the pink tax varies as a function of population characteristics. We collected county-level data on the percent of the population living in an urbanized area, median income, female population share, the share of women that are employed, and the share of women with a college degree from the Census Bureau's 2006-2010 American Community Survey (ACS). We think of median income and share of women that are employed as factors associated with female consumers' economic conditions. The female population share proxies for the relative demand for women's vs men's products in a given region, and the share of women with a college degree and share of women that are employed proxy for social norms related to preference for gendered products. Table 3 reports summary statistics of these variables.

Table 3: Demographic Variable Summary Statistics

Variable	Min	Median	Mean	Max	SD
Percent Urbanized Area	0.00	0.88	0.69	1.00	0.37
Median HH Income	19,351.00	50,014.00	52,579.00	115,574.00	13,290.91
Population Share Female	0.30	0.51	0.51	0.57	0.01
Share Women College Grads	0.07	0.34	0.35	0.73	0.09
Share Women Employed	0.25	0.53	0.52	0.72	0.05

Notes: This table provides county-level summary statistics from the 2010 American Community Survey.

3 Measuring Price Disparities

We estimate the gender price gap as the difference in average price between products targeted at men and women across retail outlets from 2016-2018. In particular, we aggregate the raw Nielsen data to the product-store-year level and model the price of product j sold by retail outlet s in year t, p_{jst}, as a function of its intended gender target, year fixed effects

Electronic copy available at: https://ssrn.com/abstract=3882214

(Γ_t) and retail outlet fixed effects (Ω_s). These fixed effects capture determinants of price that vary across location or over time and are important to the extent that different types of stores offer a larger or smaller assortment of men's and women's products. Our specification is

$$p_{jst} = \beta \cdot women_j + \Gamma_t + \Omega_s + \varepsilon_{jst}, \tag{1}$$

where the object of interest, β, is the coefficient on an indicator for whether product j is targeted at women, $women_j$. This coefficient captures the national price gap. We estimate equation (1) separately for each personal care product module.[13]

We find women face *lower* shelf prices in a majority of personal care product modules. Specifically, we estimate equation (1) with shelf price as the dependent variable. Regressions of shelf price are weighted by the number of weeks a product was available in a given store and year in order to mirror the assortment available to consumers. Results are displayed in column (1) of Table 4. Bar soap, body wash (soap - specialty), and deodorant products targeted at women are relatively more expensive than their counterparts targeted toward men, but hair coloring, razors, shampoos, and shaving creams targeted at women are relatively less expensive. Because women's and men's products may differ in size, column (2) reports differences in unit price (i.e. price per ounce or count, depending on the category). The price gaps in deodorant, bar soap, and body wash are exacerbated by the fact that women's products are typically smaller than men's products. Additionally, while product shelf prices for women's razor blades are lower, women face higher prices per individual cartridge.

We next consider the difference in the price paid for men's and women's products. This measure gives a sense of how differences in offered prices translate to differences in expenditures, which can give a sense for whether price differences create an economically significant burden for female consumers. This analysis is novel in the pink tax literature because our data reveal what people buy, in addition to what stores charge. Of course, we must keep in mind that higher prices paid do not per se indicate a higher burden if they are offset by higher quality products. We also recognize that the impact of high shelf prices for seldom-purchased products could still be large if women prefer those products but cannot afford them. These two possibilities are explored further below. Column (3) of Table 4 shows the difference in the average price paid for women's and men's products, and column (4) reports the results for unit price. Regressions are weighted by quantity sold of a product in a given store and year. Women pay higher prices for deodorants, bar soap, and body wash on both a per product and per unit basis, although the magnitude of the differences is smaller. They

[13] We restrict our analysis to product modules that are not dominated by a single gender. Specifically, we focus on modules in which neither gender accounts for more than 90% of gendered volume sales.

Electronic copy available at: https://ssrn.com/abstract=3882214

Table 4: Price Gap by Category, 2016-2018

	(1)	(2)	(3)	(4)
Category	Product Shelf Price	Unit Shelf Price	Product Price Paid	Unit Price Paid
Soap - Bar	0.41***	0.25***	0.39***	0.23***
	(0.00)	(0.00)	(0.00)	(0.00)
	4.32	0.23	4.13	0.22
	9.4%	108.1%	9.6%	101.8%
Soap - Specialty	0.72***	0.19***	0.47***	0.15***
	(0.00)	(0.00)	(0.00)	(0.00)
	4.50	0.28	4.27	0.26
	16.0%	69.4%	11.0%	55.7%
Deodorants - Personal	0.29***	0.49***	0.16***	0.43***
	(0.00)	(0.00)	(0.00)	(0.00)
	4.60	1.51	4.42	1.44
	6.4%	32.8%	3.5%	30.0%
Hair Coloring	-0.79***	-0.44***	-1.03***	-0.60***
	(0.00)	(0.00)	(0.00)	(0.00)
	8.59	8.29	8.45	8.15
	-9.2%	-5.3%	-12.1%	-7.3%
Razor Blades	-4.20***	0.68***	-3.59***	0.59***
	(0.01)	(0.00)	(0.01)	(0.00)
	21.86	3.61	21.48	3.54
	-19.2%	18.8%	-16.7%	16.8%
Razors Disposable	-0.64***	-0.18***	-0.66***	-0.17***
	(0.00)	(0.00)	(0.00)	(0.00)
	7.61	2.29	7.34	2.21
	-8.4%	-7.9%	-9.0%	-7.7%
Razors Non-Disposable	-1.08***	-1.08***	-0.97***	-0.97***
	(0.00)	(0.00)	(0.00)	(0.00)
	11.45	11.45	11.19	11.19
	-9.4%	-9.4%	-8.6%	-8.7%
Shampoo				
-Aerosol/Liq/Lotion/Powder	-1.18***	0.00***	-1.11***	0.02***
	(0.00)	(0.00)	(0.00)	(0.00)
	6.25	0.49	6.09	0.48
	-18.9%	-0.3%	-18.1%	3.2%
-Bars/Concentrates/Creams	-9.91***	-0.41***	-9.47***	-0.40***
	(0.01)	(0.00)	(0.01)	(0.00)
	17.25	1.09	17.03	1.08
	-57.4%	-37.8%	-55.6%	-36.8%
Shaving Cream	-0.42***	-0.11***	-0.25***	-0.08***
	(0.00)	(0.00)	(0.00)	(0.00)
	3.70	0.62	3.59	0.60
	-11.4%	-17.3%	-7.0%	-12.5%

Notes: For each category, the first row reports the average price gap and the second row reports the standard errors in parentheses (clustered at the store and year level). The third row reports the average price of men's products. The fourth row reports the percentage price gap, calculated as the ratio of row one over row three. Regressions are estimated separately by product module and include store and year fixed effects. *** p < .01, ** p < .05, * p < .1.

Electronic copy available at: https://ssrn.com/abstract=3882214

pay lower product prices for the remaining product modules.

Taken together, the estimates in Table 4 do not support the hypothesis that women consistently face and pay higher prices. Rather, they tell a more complex story. There are economically and statistically meaningful price differences across genders, but the direction of these differences varies across product categories. These findings contrast markedly with popular press reporting on the pink tax and highlight the importance of leveraging scanner data to provide systematic evidence on pricing across a wide array of products and retailers. In particular, our results depart from those of Bessendorf (2015)'s highly publicized survey of prices at NYC retailers. Our samples differ in three important ways. First, we utilize data for stores across the nation, not just New York City. To understand whether the difference in our conclusions represents a peculiarity of pricing in New York City, we repeat our analysis restricting to NYC retailers. Web appendix Table A14 presents the results, which are similar to our national pink gap estimates. Second, we include grocery, mass merchandiser, drug, and convenience stores, whereas Bessendorf (2015) only surveys prices of personal care products in drug stores. We will show in Table 6 that the price gap for deodorant is larger in drug stores than other channels. Finally, our analysis includes all products targeted at men and women, while Bessendorf (2015) focuses on products of the same brand targeted at different genders. In Web Appendix Table A5, we show that of the top 10 deodorant brands (which account for 88% of 2016-2018 market share) only half produce products targeted at both genders. Thus, limiting analysis to prices within brand ignores the fact that a wider assortment of products is available to consumers of both genders. Overall, the comparison of our findings highlights the need for generalizable results that can speak to the pervasiveness of the pink tax across geographies and retail channels, and it highlights the need to look at all products within a category, not just those that share a brand name.

Our aim in the rest of the paper is to help interpret the price differences documented in Table 4 by using deodorant products as a case study. We focus on deodorants because they exhibit the type of price gap that motivates recent policy–women's products are on average more expensive.

3.1 Deodorants

We begin by estimating the price gap for deodorant products from 2006-2018, the full length of the Nielsen scanner dataset. Table 5 shows that the price gap across all four price variables is smaller in this sample, both in dollar and percentage terms. Shelf prices of deodorant targeted at women are 4.6% higher than those targeted at men, and the price gap is magnified by the fact that women's deodorants are smaller in size. However, column (3) shows that women, on average, do not pay more per deodorant product than men over this longer time period. Combined, columns (1) and (3) indicate that there are more high-priced

Electronic copy available at: https://ssrn.com/abstract=3882214

Table 5: National Average Price Gap **for Deodorants, 2006-2018**

	(1)	(2)	(3)	(4)
Dependent Variable	Product Shelf Price	Unit Shelf Price	Product Price Paid	Unit Price Paid
Women	0.18***	0.38***	-0.01***	0.34***
	(0.00)	(0.00)	(0.00)	(0.00)
Avg DV for Men's Products	4.05	1.38	3.87	1.32
Percent Difference	4.6%	27.9%	-0.2%	25.8%
Observations	113,010,830	113,010,830	113,010,830	113,010,830
Stores	55,941	55,941	55,941	55,941
Years	13	13	13	13

Notes: Regressions include store and year fixed effects. Standard errors are clustered at the store and year level and reported in parentheses. *** p < .01, ** p < .05, * p < .1. Price paid regressions use quantity sold as regression weights. Shelf price regressions use number of weeks with non-missing price as regression weights. Not all stores are observed for the full sample period.

women's items than men's items, but women's purchases are relatively more concentrated on low-price offerings relative to men's purchases. Women still pay more per ounce because women's deodorants are smaller.

To understand when and where this gender price gap looms largest, we explore heterogeneity in the gender price gap as a function of store characteristics and time. To this end, we estimate variations on the regression

$$p_{sjt} = (women_j \times \mathbf{X_{st}})'\beta + \Gamma_t + \Omega_s + \varepsilon_{sjt}, \tag{2}$$

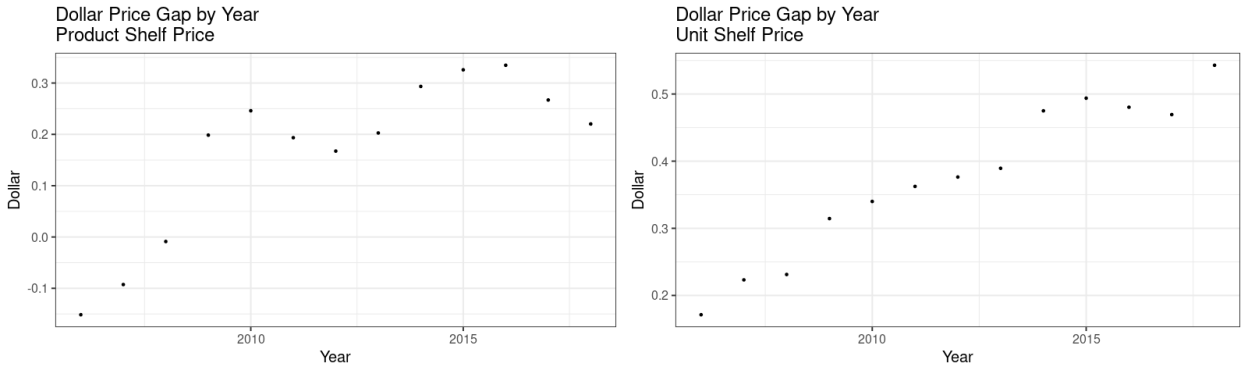
where $\mathbf{X}_{st}$ is a vector containing store characteristics or year indicators. β is the vector of coefficients of interest and summarizes heterogeneity in the price gap,

First, we analyze how the price gap in deodorants has evolved over time. Discussion of the pink tax by the press and policymakers has grown over our sample period. This heightened awareness could be because price differences have increased and become easier to notice. Conversely, as awareness of and policy surrounding this pricing practice has increased, firms may have reduced the extent to which they charge differential prices if they were concerned about negative attention amongst consumers or in the press. To explore which pattern prevails, we estimate equation (2), where $\mathbf{X}_{st}$ is a vector of year indicators and the dependent variable is product or unit shelf price. Figure 2 plots the yearly estimates and shows that the gap in deodorant prices has been trending up over time. This explains why the price gap is smaller in this longer panel relative to the shorter 2016-2018 panel. We note a slight decrease in the product price gap beginning in 2016, but the unit price gap has remained high, suggesting that women's products may have become cheaper and smaller or

Electronic copy available at: https://ssrn.com/abstract=3882214

men's products became larger and more expensive in recent years.

Figure 2: Price Gap Over Time



Notes: The left panel plots the gap in price per product. The right panel plots the gap in unit price (price per oz).

We next describe heterogeneity in the magnitude of gender price disparities across retail formats. We test whether the relative prices of men's and women's deodorants vary across food, drug, mass merchant, and convenience stores. Our results indicate that the relative shelf price difference between men's and women's deodorants is largest at grocery and drug stores and smallest at convenience stores. In addition, women are much more likely to buy low-priced items at drugstores compared to men. On a per ounce basis, women face higher prices and pay more than men at all four retail formats. This analysis highlights how our study builds on prior analysis from Bessendorf (2015). We look not only across channels, where there are meaningful differences in the price gap, but also at prices paid.

We also test whether the gender price gap is correlated with the demographics in a store's county: the percent of the population living in an urbanized area, the median income, the female population share, the share of women that are employed, and the share of women with a college degree. These demographic variables are intended to capture factors related to the economic conditions of female consumers, as well as factors that may be correlated with gender norms and preferences for gendered products. Each of the demographic variables is normalized using the formula $z_s = \frac{x_s - \bar{x}}{\sigma_x}$ (see Table 3 for the related summary statistics). Thus, in Table 7, the coefficient on $Women$ represents the price gap in a county with average demographics, and each interaction coefficient can be interpreted as the change associated with a one standard deviation increase in the corresponding demographic variable. We find that the price gap is larger in more urban areas, higher income areas, and areas with a larger share of female college graduates. The gender price gap is lower in areas in which women comprise a larger share of the population.

Electronic copy available at: https://ssrn.com/abstract=3882214

Table 6: Price Gap by Retail Format

Dependent Variable	(1) Product Shelf Price	(2) Unit Shelf Price	(3) Product Price Paid	(4) Unit Price Paid
Women X Conv	-0.60***	0.09***	-0.49***	0.15***
	(0.02)	(0.01)	(0.02)	(0.01)
Women X Drug	0.25***	0.45***	-0.06***	0.39***
	(0.00)	(0.00)	(0.00)	(0.00)
Women X Food	0.28***	0.39***	0.14***	0.34***
	(0.00)	(0.00)	(0.00)	(0.00)
Women X Mass	-0.03***	0.28***	-0.11***	0.31***
	(0.00)	(0.00)	(0.00)	(0.00)
Observations	113,010,830	113,010,830	113,010,830	113,010,830
Stores	55,941	55,941	55,941	55,941
Years	13	13	13	13

Notes: Regressions include store and year fixed effects. Standard errors are clustered at the store and year level and reported in parentheses. *** p $<$ .01, ** p $<$ .05, * p $<$ .1. Price paid regressions use quantity sold as regression weights. Shelf price regressions use number of weeks with non-missing price as regression weights. Not all stores are observed for the full sample period.

Table 7: Price Gap as a Function of County Demographics

Dependent Variable	(1) Product Shelf Price	(2) Unit Shelf Price	(3) Product Price Paid	(4) Unit Price Paid
Women	0.17***	0.38***	-0.03***	0.32***
	(0.00)	(0.00)	(0.00)	(0.00)
Women X % Urban Area	0.03***	0.02***	0.00	0.02***
	(0.00)	(0.00)	(0.00)	(0.00)
Women X Median Income	0.03***	0.01***	0.04***	0.02***
	(0.00)	(0.00)	(0.00)	(0.00)
Women X Pop Share Female	-0.02***	-0.01***	-0.04***	-0.03***
	(0.00)	(0.00)	(0.00)	(0.00)
Women X Share Female College	0.04***	0.03***	0.04***	0.05***
	(0.00)	(0.00)	(0.00)	(0.00)
Women X Share Female Employed	0.00***	0.00***	0.02***	-0.01***
	(0.00)	(0.00)	(0.00)	(0.00)
Observations	113,010,830	113,010,830	113,010,830	113,010,830
Stores	55,941	55,941	55,941	55,941
Years	13	13	13	13

Notes: Regressions include store and year fixed effects. Standard errors are clustered at the store and year level and reported in parentheses. *** p $<$ .01, ** p $<$ .05, * p $<$ .1. Price paid regressions use quantity sold as regression weights. Shelf price regressions use number of weeks with non-missing price as regression weights. Not all stores are observed for the full sample period. Demographic variables are normalized using $z_s = \frac{x_s - \bar{x}}{\sigma_x}$. Thus, the coefficient on $Women$ represents the pink tax in a county with average demographics, and each interaction coefficient can be interpreted as the change associated with a one standard deviation increase in the corresponding demographic variable.

Electronic copy available at: https://ssrn.com/abstract=3882214

We close this section by investigating whether the price gaps for deodorant that we document in Table 5 reflect differences in regular prices or differences in promotional behavior. We define a price promotion as a week when a product (UPC) is less than 95% of its regular price (see Web Appendix A for details on the algorithm we use to construct the regular price series). For each store-product-year combination, we compute the average regular (un-promoted) price, the fraction of weeks with a price promotion, and the average percent discount, conditional on a promotion occuring. We then re-estimate equation (1) using each of these measures as the dependent variable.

Estimates are presented in Table 8. In terms of regular price, women's items are on average 16 cents or 4% more expensive relative to an average men's price of $4.14. This is similar to the documented price gap in price charged shown in column (1) of Table 5. Turning to promotions, women's items are discounted 4% less often and conditional on a discount, the average percent discount is 7% smaller for women's items. Together, these results suggest that the documented discrepancies in price charged are driven by both differences in regular prices, as well as differences in the prevalence and depth of discounts.

Table 8: Regular Prices and Discounting Behavior for Men's and Women's Deodorant Products

Dependent Variable	(1) Regular Price	(2) Frac Weeks w/ Discount	(3) Discount Percent
Women	0.16***	-0.004***	-0.017***
	(0.001)	(0.0001)	(0.0001)
Store FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Avg DV for Men's Products	4.14	0.095	0.23
Percent Difference	3.9%	-4.2%	-7.2%
Observations	104,498,780	104,498,780	68,498,629
Stores	55,954	55,954	55,954
Years	13	13	13

Notes: Standard errors are clustered at the store and year level and reported in parentheses. *** p $<$ .01, ** p $<$ .05, * p $<$ .1. Not all stores are observed for the full sample period.

4 Costs

In the simplest model of firm pricing, two factors could contribute to the price disparities we document in Section 3: differences in manufacturing costs across products targeted toward men vs women; and differences in demand elasticities for products targeted toward men vs women. Identifying whether elasticities or costs drive the gender price gap is important for understanding the scope for and implications of policy proposals. For example, if costs are the same but elasticities differ, then mandating price parity across gendered products could increase total surplus. However, the same policy could reduce total surplus if the true source of price differences is manufacturing costs.

Electronic copy available at: https://ssrn.com/abstract=3882214

Wholesale Prices

One possibility is that the retail gender price gap that we document above simply reflects upstream price disparities between women's and men's products. Indeed, if women's products cost more to produce, then we would expect manufacturers and wholesalers to charge higher prices for women's products. Similarly, provided that upstream firms can exercise market power, they ought to set higher prices for women's products if women have relatively inelastic demand. This section documents differences in the wholesale prices of men's and women's deodorant products using the PromoData Price-Trak dataset, which records wholesale prices at the UPC-DMA-year-level between 2006 to 2011. In Table 9, we employ the following regression specification where j denotes a product (UPC), m denotes a market (DMA), and t denotes a year:

$$c_{jmt} = \gamma_0 + \gamma_1 \cdot women_j + \Gamma_t + \Omega_m + \varepsilon_{jmt} \tag{3}$$

The variable $women_j$ is an indicator for a women's product, and the null hypothesis is that products targeted at men and women have similar wholesale prices ($\gamma_1 = 0$). Γ_t and Ω_m are year and market fixed effects, respectively. We estimate this regression for two measures of wholesale price c_{jmt}, list price and deal price after trade fund allowances, and we analyze both product and unit prices. Regressions are weighted using the total number of store-week combinations with non-missing price for a product in a market and year.

Table 9 presents estimates of the wholesale price gap, γ_1: wholesale prices do not mirror the retail price patterns described in Section 3. Columns (1) and (3) show that wholesale list and deal product prices are about \$0.12 or 4% lower for women's deodorant products than for men's, although estimates are statistically insignificant. On the other hand, columns (2) and (4) report estimates of γ_1 for list and deal unit prices and show that women's deodorants are \$0.16 or 16% more expensive than men's deodorants per ounce. Although women's productshave cheaper wholesale product prices, the fact that women's products are smaller makes them more expensive per unit.

We note that these estimates are not easily comparable to the estimates of retail price differences in Table 5 because the sample of products and years differ. The wholesale data is a shorter panel (2006-2011 vs 2006-2018) and covers fewer markets (29 vs 209) and products (981 vs 3,662 deodorant UPCs). Thus, to obtain comparable retail price estimates, we re-estimate the retail gender price gap in the Nielsen Scanner data, subsetting to the set of products, markets, and years for which we observe wholesale prices. We use the same regression specification as in equation (1).

Electronic copy available at: https://ssrn.com/abstract=3882214

Table 9: Wholesale Price Differences

Dependent Variable	(1) Product List Price	(2) Unit List Price	(3) Product Deal Price	(4) Unit Deal Price
Women	-0.12	0.16**	-0.1	0.16**
	(0.1)	(0.04)	(0.1)	(0.04)
Market FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Avg DV for Men's Products	2.8	0.98	2.75	0.96
Percent Difference	-4.2%	16.4%	-3.8%	16.9%
Observations	2,989	2,989	2,989	2,989
Markets	28	28	28	28
Years	6	6	6	6

Notes: Wholesale prices for products that appear in both the PromoData and the 2006-2011 Nielsen Scanner data. Standard errors clustered at the market and year level reported in parentheses. *** p $<$ .01, ** p $<$ .05, * p $<$ .1. Regressions use total number of store-weeks with non-missing price in Nielsen Scanner data as regression weights. Not all stores or products are observed for the full sample period.

Columns (1) and (5) in Table 10 present retail price disparity estimates for the PromoData subset of years, markets, and products. As before, the shelf price regressions are weighted by the number of weeks a product is available on the shelf of a given store. The results indicate a statistically insignificant and economically modest difference in product shelf prices of -$0.02. This contrasts with the results in Table 5, which showed women's products were $0.18 more expensive in the full sample. The estimate for the difference in shelf price per oz is more similar across the two samples. The unit shelf price of women's deodorant products is $0.30 higher than men's, compared to $0.38 in Table 5. Columns (2) and (6) replicate Table 9 in this dataset to make comparisons more convenient. While the estimates are noisier, the magnitudes are the same as those in Table 9. (Column (2) differs by $0.01 because of rounding.) Next, we compare the retail margins and markups for men's and women's products. Columns (3) and (7) report the difference in the dollar margin of women's and men's products. The dependent variable is the difference in the retail shelf price and the wholesale price. The estimates show that the per product margin for women's products is $0.09 higher than for men's products. Per unit, the margin on women's products is $0.14 more than on men's products. Finally, columns (4) and (8) in Table 10 show that markups are higher for women's products than men's products. These columns present estimates from the regression

$$p_{jst} = \mu_0 c_{jmt} + \mu_1 c_{jmt} \times women_j + \varepsilon_{jst} \tag{4}$$

Electronic copy available at: https://ssrn.com/abstract=3882214

Electronic copy available at: https://ssrn.com/abstract=3882214

Table 10: Retail Price Gap and Markups for Products and Years with Wholesale Price Data

Dependent Variable	(1) Product Shelf Price	(2) Product List Price	(3) Product Margin	(4) Product Shelf Price	(5) Unit Shelf Price	(6) Unit List Price	(7) Unit Margin	(8) Unit Shelf Price
Women	-0.02	-0.11	0.09		0.3***	0.16**	0.14***	
	(0.11)	(0.13)	(0.05)		(0.07)	(0.06)	(0.03)	
Wholesale Price				1.14***				1.14***
				(0.00)				(0.00)
Women X Wholesale Price				0.15***				0.25***
				(0.00)				(0.00)
Store FE	Yes	Yes	Yes	No	Yes	Yes	Yes	No
Year FE	Yes	Yes	Yes	No	Yes	Yes	Yes	No
Avg DV for Men's Products	3.96	3.06	0.9		1.39	1.07	0.32	
Percent Difference	-0.6%	-3.6%	9.6%		21.3%	14.8%	43.5%	
Observations	696,748	696,748	696,748	696,748	696,748	696,748	696,748	696,748
Markets	10,872	10,872	10,872		10,872	10,872	10,872	
Years	6	6	6		6	6	6	

Notes: Standard errors are clustered at the store and year level and reported in parentheses. *** p $<$ .01, ** p $<$ .05, * p $<$ .1. Shelf price regressions use number of weeks with non-missing price as regression weights. Not all stores are observed for the full sample period.

and $(\mu_0 - 1)$ and $(\mu_0 + \mu_1 - 1)$ can be interpreted as the retail markup for men's and women's products, respectively. The retail markup for men's products both per product and per unit is 14%. Women's products face higher retail markups of 29% per product and 49% per unit.

Taken together, the results in Tables 9 and 10 provide some evidence that retail price differences are not solely driven by upstream price differences. If differences in retail prices were due entirely to differences in manufacturing costs, we would expect to see similar differences in wholesale prices. Instead, we find that wholesale prices per unit for women's deodorant products are significantly lower than prices for men's products and retail markups are higher on women's products.

The analysis in this section comes with a caveat. The PromoData represents the prices that 12 major grocery store wholesalers charged to independent grocery stores. However, the Nielsen scanner data is primarily composed of large chain stores in the United States. These large chains likely do not purchase from wholesalers, but instead act as their own distributors and negotiate their own wholesale prices directly with the manufacturer. The price gap in wholesale prices from the PromoData can be informative about the wholesale prices faced by Nielsen retailers if the differences in prices that the two sets of retailers pay do not vary systematically by gender. While we think this assumption is reasonable, this would be violated if, for example, large chains tend to get larger quantity discounts for men's products than for women's products.

Price Disparities Controlling for Characteristics

One challenge in interpreting the average price differences that we document between men's and women's products is that products may differ along other dimensions. If deodorant attributes differ systematically across products targeted at men vs women that could lead to asymmetries in manufacturing costs, which could in turn generate a gender price gap. We have already provided evidence of modest differences in wholesale prices, which suggest that women's products are unlikely to be more costly to produce. We have also already shown that women's products are smaller, which also suggests they ought to be cheaper to produce. In this section, we provide complimentary evidence on costs by directly examining whether men's and women's products differ on non-price product attributes.

Data on attributes comes from the Nielsen product description field and from Label Insight, which covers 26% of deodorant products sold in the scanner dataset. The Nielsen product description field indicates whether each product is an antiperspirant. The Label Insight data records marketing claims made on product packaging, such as "aluminum free" or "moisturizing." Table 11 describes the incidence of these claims separately for men's and women's products. The final column calculates the difference in the likelihood of each claim

Electronic copy available at: https://ssrn.com/abstract=3882214

between genders. Women's products are more likely to be antiperspirants and moisturizing.[14] However, they are equally likely to be aluminum and cruelty free, as well as deodorizing. Overall, we detect differences in two out of seven of the attributes, which leaves open the possibility that non-price characteristics could play a role in the gender price gap.

Table 11: Deodorant Attributes for Men's & Women's Products

	% UPCs Within Gender		
Attribute	Women	Men	Diff
Antiperspirant (RMS)	91.8%	69.3%	22.5%***
Total UPCs (RMS)	2,126	2,113	
Aluminum Free	8.6%	7.3%	1.2%
Antiperspirant (LI)	63.9%	55.7%	8.2%***
Cruelty Free	3.1%	2.9%	0.2%
Deodorize	71.2%	69.1%	2.1%
Longlasting	26.6%	30.0%	-3.4%
Made In USA	17.1%	15.0%	2.1%
Moisturizing	8.6%	1.2%	7.3%***
Total UPCs (LI)	549	560	

Notes: This table summarizes differences in product attributes and marketing claims for men's and women's deodorant products. The first row reports the share of products that are antiperspirants, per the Nielsen RMS product description. Subsequent rows report the incidence of different marketing claims in the Label Insight data.

To probe the role of attributes further, we augment specification (1) to control for the characteristics in Table 11. Estimates are presented in Table 12. Column (1) reports the price difference for the sample of Label Insight products following our baseline specification (1). Before controlling for attributes, products targeted at women are are on average $0.44 more expensive. This gender price gap shrinks to $0.29 when we control for product characteristics (column (2)). Thus, approximately 35% of the gender price gap can be explained by differences in observed attributes. The same is true when looking at the price per ounce in columns (3) and (4).

[14]The baseline prevalence of antiperspirants is higher in the full sample relative to the measure from the Label Insight data. This could be because the products only observed in the RMS data are more likely to be antiperspirants. Alternatively, this could indicate that some products have attributes that are not explicitly mentioned on the packaging. In other words, the absence of a marketing claim does not necessarily imply the absence of a product characteristic.

Electronic copy available at: https://ssrn.com/abstract=3882214

Table 12: Price Differences Conditional on Attributes

	(1)	(2)	(3)	(4)
DV	Unit Shelf Price	Unit Shelf Price	Unit Price Paid	Unit Price Paid
Women	0.44***	0.29***	0.36***	0.25***
	(0.001)	(0.001)	(0.001)	(0.001)
Attributes	N	Y	N	Y
Avg Men's DV	1.40	1.40	1.21	1.21
% Difference	31.6%	20.7%	30.0%	20.7%

We highlight that moving beyond these price differences to pinpointing a "pink tax" is still challenging because firms select gender targeting and other attributes simultaneously. One concern is that our data feature marketing claims that ostensibly reflect demand for different characteristics and not simply their presence. In other words, men's and women's products could be equally moisturizing, but if women prefer moisturizing products, the claim may be highlighted disproportionately on packaging targeted at women. Further, even if the prevalence of an attribute truly differs across men's vs women's items, associated price differences are not necessarily due to cost differences. For example, if women have a higher willingness to pay for moisturizing, these demand-side differences could lead firms to set higher prices for women's products with moisturizing relative to men's products without moisturizing. Namely, any observed differences in attributes need not imply that the associated price differences are due to differences in manufacturing costs. Differential demand for attributes that are relatively cheap to produce could generate a price gap that is due to differences in preferences rather than differences in costs. Thus, absent data on manufacturing costs, we next turn to estimating demand elasticities.

5 Elasticities

The classic economic rationale for price discrimination is demand heterogeneity. In the context of the pink tax, if demand for women's deodorants is less elastic than for men's, a profit-maximizing retailer should impose higher markups on women's deodorants. To investigate this mechanism, we estimate a log-log model of demand for men's and women's deodorant, where the quantity sold of brand j in store s in week t is specified as:

$$\log(q_{jst}+1) = \alpha_{js} + \sum_{k \in \mathcal{J}_m} \beta_{jkm} \log(p_{kst}) + \tau_{jm(s)t} + \epsilon_{jst}. \tag{5}$$

The main parameter of interest is β_{jjm}, the coefficient on product j's own price, p_{jst}. We also include the prices of other deodorant products as controls, which is captured by the sum over products $k \in \mathcal{J}_{m(s)}$, where $\mathcal{J}_{m(s)}$ denotes the set of included brands in the market

Electronic copy available at: https://ssrn.com/abstract=3882214

$m(s)$ where store s is located. Thus, parameters β_{jkm} are cross-price elasticities of brand j's demand with respect to the price of brand k, averaged across stores in market m. The additional control variables include α_{js}, a set of brand-store fixed effects that accounts for different levels of demand across stores, and $\tau_{jm(s)t}$, a set of brand-market-time fixed effects that controls for market-level demand shocks that may vary over time. Hitsch, Hortaçsu, and Lin (2021) employs a similar demand specification and, like them, our preferred specification operationalizes the model using brand-county-month fixed effects.[15][16]

When estimating equation (5), the identifying assumption is that the included fixed effects account for any unobservables that may be correlated with prices and volume sales such that, net of the fixed effects, the remaining variation in prices is as good as random. One nice feature of the deodorant category is that it exhibits relatively predictable seasonality in demand and is less likely to experience transient demand shocks compared to a category like beer. The model's county-specific month fixed effects will capture demand shocks due to, for example, geography-specific seasonal weather patterns. Thus, any endogeneity concerns must stem from store-specific deviations from county demand patterns. We think the most likely threat to our identification strategy is other store-specific marketing activity like feature or display advertising that might be coordinated with pricing and could directly affect sales. If we fail to account for this store-specific marketing activity in our analysis, our elasticity estimates could be biased. Because we are primarily interested in the relative elasticities of demand across genders and the resulting implications for gender-based pricing, it would be especially problematic if products targeted to one gender were more likely to be supported with other marketing activity and thus any bias were asymmetric across genders. To explore this possibility, we first collect and summarize data on feature and display advertising for the subset of stores (17%) for which this data is available. Deodorant products are not heavily supported with feature and display advertising. An average deodorant product is on feature (display) in 7% (1%) of weeks in a year. Comparing across genders, women's products are on feature slightly less often than men's, but they are on display slightly more frequently than men's. We then probe the sensitivity of our results by re-estimating equation (5) including feature and display variables. The results of this robustness analysis are discussed below.

We estimate demand using specification (5) for the top 5 brands for each gender in each

[15] We also estimate a specification with brand-county-week fixed effects. While the level of elasticity changes slightly, the pattern across genders is consistent. As another robustness, we emulate DellaVigna and Gentzkow (2020) and estimate the log-log model with store-year and week-of-year fixed effects, both with and without instruments measuring the average price of the focal product in stores of the same chain outside the focal store's county. However, our preferred specification is the one with county-month fixed effects because these fixed effects account for local weather patterns that may be important drivers of demand for deodorant products.

[16] Our specification deviates from Hitsch, Hortaçsu, and Lin (2021)'s in two ways. First, we estimate an average store-level elasticity in each market, as opposed to estimating the full distribution of store-level elasticities. We also omit indicators for "deal" promotions in our specification because including the deal indicators produces improbably inelastic estimates. However, the asymmetries in elasticities by gender that we document below are robust to the inclusion of deal indicators.

Electronic copy available at: https://ssrn.com/abstract=3882214

state from 2014-2018.[17] For each week and store in the scanner data, we average the shelf price and sum quantity across all UPCs associated with a given brand description and gender categorization. For example, we aggregate over all *Degree* men's deodorant UPCs in each week and store. Aggregating to the brand level and subsetting to the most popular products reduces the computational burden of the model. We aggregate the UPC-level data on feature and display to the brand level by computing the fraction of a brand's UPCs that were on feature or display in a given week. Following Hitsch, Hortaçsu, and Lin (2021), who employ a similar demand specification, we remove stores in each county that have more than 20% of weeks with an imputed shelf price. We allow the set of products included in the demand estimation to vary across markets. Table A7 in Web Appendix H lists the brand-gender pairs included in our sample and the number of markets in which each appears. For computational reasons, we estimate the model separately for each state/brand-gender pair and obtain for each state/brand-gender the average elasticity across stores in that state.

The resulting distribution of own price elasticity estimates is summarized in Table 13. We find that demand for women's deodorant products is less elastic than demand for men's deodorant products. The mean elasticity of a men's deodorant product is -1.56 compared to -1.23 for a women's deodorant product, and the distributions of elasticities follow the same pattern. The difference in average elasticities across genders is statistically significant ($p < 0.01$). Furthermore, including the additional controls for feature and display advertising changes the level of the estimated elasticities somewhat, but the pattern that men's products have more elastic demand than women's products is robust.

The differences in elasticities across genders are not driven by one outlier brand. Figure 3 shows violin plots of the distribution of market-level elasticities for the top 5 brands of each gender by national sales. The vertical line shows the median elasticity and the dot shows the mean elasticity of each brand-gender. Consistent with the pattern in Table 13, demand

[17] The market share of the top 5 brands of each gender in the median market is about 70%.

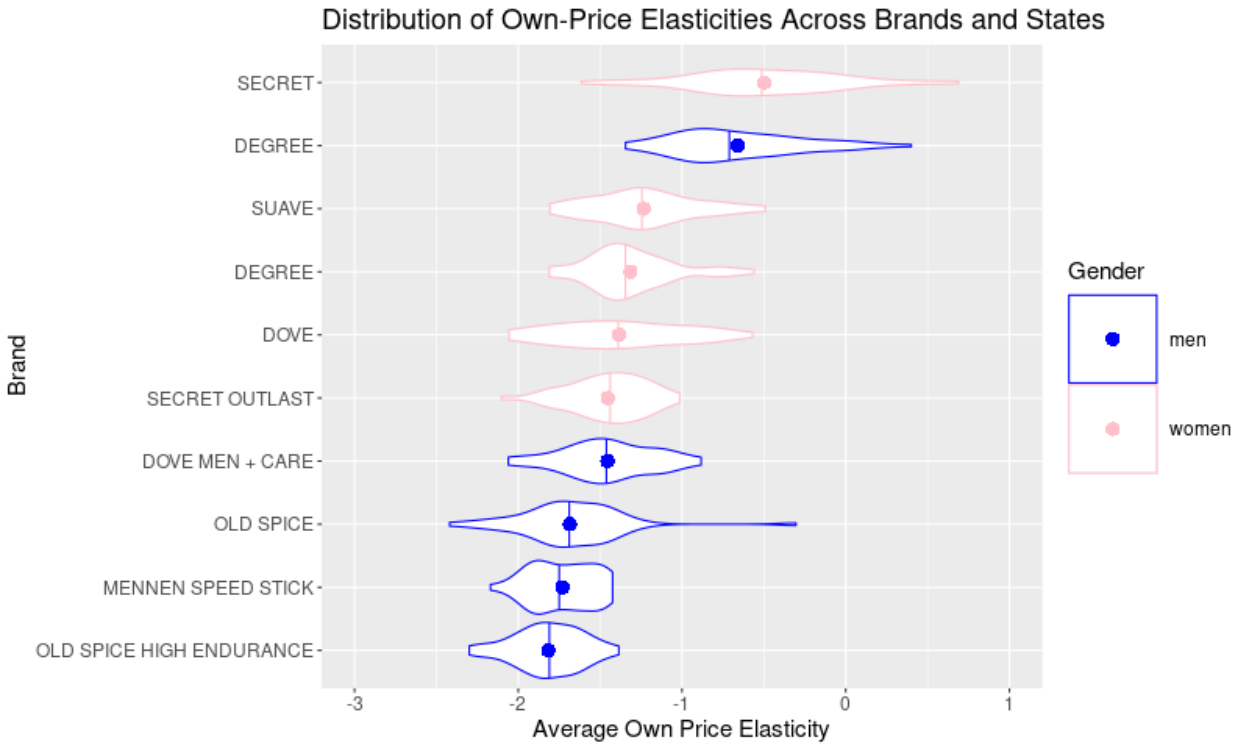
Table 13: Own Price Elasticity Estimates by Gender, 2014-2018

Specification	Gender	Mean	25th Percentile	Median	75th Percentile
Store & County-Month FE	men	-1.47	-1.83	-1.60	-1.26
	women	-1.17	-1.50	-1.28	-0.89
Store & County-Month FE	men	-1.55	-1.86	-1.48	-1.17
+ Feature & Display	women	-1.29	-1.63	-1.26	-0.88

Note: This table summarizes the distribution of estimated elasticities across brand-genders and states.

Electronic copy available at: https://ssrn.com/abstract=3882214

Figure 3: Own-Price Elasticities of Demand, 2014-2018



Note: Violin plots of the distribution across states of own-price elasticities for the top 5 brands of each gender by national sales. The vertical line shows the median elasticity and the dot shows the mean elasticity of each brand-gender.

for women's brands is generally less elastic than for men's brands. *Degree* men's deodorant is an exception with a mean and median elasticity that is less elastic than four of the five women's brands.

Thus far we have focused on comparing the magnitude of elasticities across genders. Turning to the baseline level of elasticities, we note that the estimates are relatively inelastic, implying high markups. For comparison, Hitsch, Hortaçsu, and Lin (2021) estimate a median brand-level elasticity of -1.93 across 2,000 CPG products (the inner-quartile range of their estimates is [-3.22, -0.76]).

The cross price elasticity estimates, which we omit for brevity, are noisy and difficult to interpret. For example, most cross price elasticity estimates are insignificant, implying minimal price competition in the category. This is perhaps not surprising because cross price elasticities are notoriously difficult to estimate in the log-log demand model.

Notwithstanding, our finding that demand for women's deodorant products tends to be more inelastic compared to demand for men's deodorant products suggests that the price gap for deodorants that we documented in Section 3 may be at least partially driven by differential markups. This leaves open the possibility that there could be scope for price-

Electronic copy available at: https://ssrn.com/abstract=3882214

parity regulation to increase consumer welfare. An important next step is to unpack the drivers of the differences in elasticities we documented here.

6 What Drives Differences in Elasticities?

This section explores why demand for women's deodorant products is relatively more inelastic. Such an understanding can help inform whether market intervention would likely lead to an increase in consumer welfare. We consider several potential explanations for the documented differences in demand elasticities and conduct empirical analyses to probe each one.

Assortments

Our elasticity estimates measure residual demand for a particular product at a particular retailer–that is, we measure elasticities for men's and women's products, rather than elasticities for male and female consumers. Thus, differences in assortments might engender differences in elasticities even if men and women share similar underlying price sensitivities. We therefore compare assortments for men's and women's products. Table 14 describes product- and brand-level assortments among RMS retailers. Across all formats, the average brand assortment size is larger for men's vs. women's deodorant products. This is echoed in product assortment sizes, with the exception of mass merchandizers.

This pattern also speaks to economies of scale. A similar level of demand spread across a smaller assortment should lead to higher unit sales per product. This is confirmed by the finding that, per year and summing across all stores, the average women's UPC has 60% higher unit sales than the average men's UPC.[18] Any scale economies in deodorant manufacturing should lead to relatively lower prices for women's products, another piece of evidence against a cost-based explanation.

[18]The average men's UPC sells 110,516 units in the Scanner data per year, while the average women's UPC sells 177,088 units.

Electronic copy available at: https://ssrn.com/abstract=3882214

Table 14: Men's and Women's Deodorant Assortments

	Men	Women	Difference
# Products Overall	2,113	2,126	–13
# Products per Store:			
Drugstore	159	140.7	18.318***
Food	144.5	123.1	21.44***
Mass Merch	90.3	93.9	–3.59***
Convenience	2.4	1.6	0.833*
# Brands Overall	158	114	44
# Brands per Store:			
Drugstore	33.57	25.7	7.87 ***
Food	29.66	22.03	7.63***
Mass Merch	22.5	18.91	3.59***
Convenience	1.73	1.25	0.48***

Notes: A product is a UPC. We estimate the difference across genders in the average number of products and brands per store via a regression of store-year-gender assortment sizes on an indicator for gender interacted with dummies for the four retail formats.

Advertising

This section explores whether advertising is linked to the documented differences in elasticities and consequently shelf prices for men's and women's deodorant products. We begin by exploring TV advertising volumes using data from Ad Intel. The data includes both network and spot advertisements that aired between 2010 and 2018 in the US. Ad spots are identified with a brand variant (e.g. "Old Spice Fresh Collection Antiperspirant"), which is finer than a parent brand (e.g. "Old Spice") but coarser than a UPC. We determine gender targeting for brands based on the categorization of UPCs within that brand and on the use of gendered words in the brand variant title and the creative description. Web Appendix D provides more detail on the procedure. In total, of the forty-two brand variants advertised on television, 18 are targeted at women and 16 at men. Figure 4 plots the volume of ads on network and spot TV by gender over time–it is apparent that the lion's share of deodorant ad occurrences promote women's products, particularly in the spot market. A t-test that ads are equally likely to be aimed at men and women rejects with a t-statistic of 133 for network ads and 621 for spot ads.

Electronic copy available at: https://ssrn.com/abstract=3882214

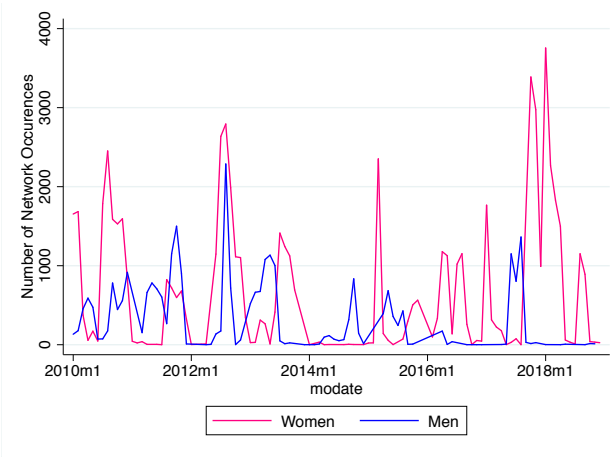
Table 15: Ad Elasticities for Deodorant

Brand	Estimated Ad Elasticity
Degree	0.1104
	(0.0316)
Dove	0.0491
	(0.0151)
Secret	0.0166
	(0.0085)

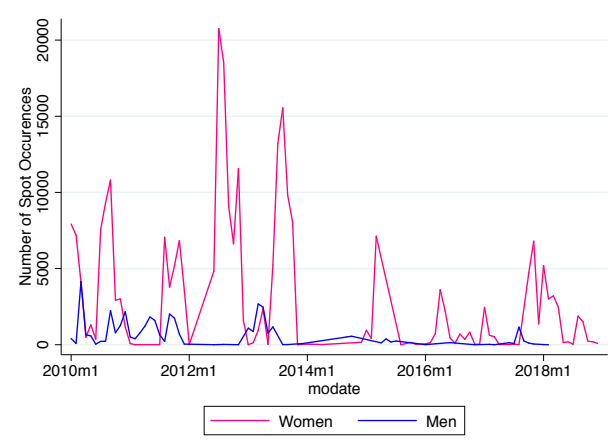
Notes: Advertising elasticities from B. Shapiro, Hitsch, and Tuchman (Forthcoming) for personal care products.

Figure 4: Ad Occurrences over Time by Gender

(a) Network



(b) Spot



Higher advertising for women's products could generate price disparities through two avenues. First, advertising could increase perceived product differentiation, thereby affecting demand elasticities directly (see Bagwell (2007) for a longer discussion of this mechanism). However, estimates from B. Shapiro, Hitsch, and Tuchman (Forthcoming) suggest that advertising elasticities are small for established CPG products overall. Interestingly, as shown in Table 15, the three deodorant brands in their sample all have estimated ad elasticities above the median elasticity of 0.014 in their sample of 288 brands.

Advertising might also affect pricing if firms consider advertising expenditures as costs relevant for pricing decisions. For example, if manufacturers price based on average cost, then increased advertising expenditures will be passed on to female consumers. To examine this possibility, we plan to estimate specification (1) controlling for GRPs.

Electronic copy available at: https://ssrn.com/abstract=3882214

Brand Loyalty

Holding price sensitivity fixed, demand for women's products may be relatively inelastic if women have stronger preferences over deodorant products and therefore are less likely to switch to a lower-priced alternative. That is, perhaps women are more brand-loyal. We test this hypothesis by examining the shopping behavior of one-person households from 2006-2018. In particular, we test whether women are less likely to switch brands. For each person, we count the number of distinct UPCs, brands, and parent brands they purchase, and we also compute the concentration of their purchases across parent brands, a brand "HHI." These descriptives are presented in Table 16. Men and women show similar switching patterns: they purchase a similar number of products and brands, but women purchase a wider variety of parent brands.[19] One challenge in interpreting these patterns is that purchase decisions reflect both prices and preferences. We hope to probe further using a characteristics-based demand model.

Table 16: Brand Loyalty Descriptives

Variable	Men	Women	Diff
Unique UPCs	5.70	5.56	0.14*
	(6.78)	(6.40)	
Unique Brands	3.24	3.26	-0.03
	(3.07)	(2.98)	
Unique Parent Brands	2.55	2.70	-0.16***
	(2.03)	(2.11)	
Parent Brand HHI	0.63	0.60	0.02***
	(0.31)	(0.31)	

Notes: Data comprises deodorant purchases by single-person households in the Nielsen Homescan panel 2006-2018. Standard deviations in parentheses. Difference in means *** $p < .01$, ** $p < .05$, * $p < .1$.

7 Discussion

Gender inequalities in the labor market have spawned a substantial literature in economics and a spate of federal and state regulations, such as the Equal Pay Act of 1963. In this paper,

[19] Panelists may switch for reasons outside of their control, like a product getting discontinued. On average, female panelists are observed for fewer years in the data than male panelists, so male panelists may be more affected by discontinuations. However, this does not drive the results. The patterns are the same if we summarize the number of products and brands per household and year, rather than just household.

Electronic copy available at: https://ssrn.com/abstract=3882214

we show that the recent concern that such gender price discrimination extends to personal care products is unfounded. Our analysis of prices and sales at a panel of 39,697 national grocery, convenience, drug, and mass merchandise stores does not support the existence of a price premium for women's products. To the contrary, we find that men's products command higher prices in seven of the ten categories that we study.

Whether these simple price differences speak directly to a pink tax is not entirely clear. One concern is that comparing prices across products targeted at different genders is not an apples-to-apples comparison if non-price attributes differ systematically across genders. Relatedly, the price gaps that we do observe might reflect differences in manufacturing costs rather than differences in markups. To explore these possibilities, we focus on deodorants, a category where women's products are indeed more expensive. We find that the gender price gap falls by 30% when we control for observable attributes, but that wholesale costs do not explain the remaining price differences. Rather, our results suggest that the price gap reflects differences in elasticities for men's and women's products. Thus, a profit-maximizing firm ought to charge a higher markup on women's products in this category.

An important ingredient for this type of price discrimination is the low rate of abitrage in personal care; we find that women are significantly more likely to buy women's products and men to buy men's products. If men and women essentially comprise separate markets, then we might think of this type of pricing as third degree price discrimination, in which case mandating price parity across genders could increase total surplus. Such regulation is atypical in CPG but has precedent in other arenas. For example, the U.S. Fair Housing Act prohibits charging different rent to lessees or offering different mortgage interest rates to lenders based on their gender, race, or other protected classes. As another example, the EU's Unfair Commercial Practices Directive prohibits firms like airlines for adding fees for EU customers solely based on their nationality. That said, it is unclear how such regulations could be enforced for CPG products. For example, New York State law Section 391-U prohibits differential pricing for "substantially similar" goods of the same brand. However, many manufacturers sell men's and women's items under different brands (e.g. P&G owns both Old Spice and Secret). Thus, banning within brand price discrimination would not eliminate the gender price gap. Regulators may instead want to consider policies that mandate price parity within manufacturer rather than within brand and to specify price parity on a per oz instead of per product basis.

Further, the long-term benefits of price parity regulations are murky. Policies that restrict firms' ability to set prices may inhibit entry or otherwise reduce assortments, which in turn may reduce welfare. The welfare consequences then depend on the relative importance of match quality for personal care products relative to economies of scale in production, objects that are hard to identify empirically. This is particularly true in an industry where product differentiation may be spurious (C. Shapiro 1982). The crux of the issue lies in understanding

Electronic copy available at: https://ssrn.com/abstract=3882214

why women (men) exhibit such a strong preference for products targeted at women (men) that they are willing to pay more for them. One possibility is that gender cues, such as color, truly increase the utility that consumers derive from deodorants, razors, or shampoos. A second possibility is that consumers believe that color differences indicate other differences between products (e.g., perhaps there are ingredients in a men's deodorant that are ill-suited for women's use). Bronnenberg et al. (2015) provide evidence of this type of information asymmetry in the CPG context, and suggest it can be linked to advertising. Indeed, we find evidence of large asymmetries in advertising for women's deodorant products. If gender labeling is a case of spurious differentiation, then it may soften price competition, leading to higher prices for both men and women. In this case, a market intervention that corrects information asymmetries could increase competition and lead to lower prices and higher consumer welfare for both men and women. With our current data we are not able to distinguish between these stories, but this seems like a promising avenue for future research.

References

Anderson, Eric T. and Duncan I. Simester (2008). "Does Demand Fall When Customers Perceive That Prices Are Unfair? The Case of Premium Pricing for Large Sizes". In: *Marketing Science* 27.3, pp. 492–500. ISSN: 07322399, 1526548X. URL: http://www.jstor.org/stable/40057149.

Ayres, Ian (1991). "Fair Driving: Gender and Race Discrimination in Retail Car Negotiations". In: *Harvard Law Review* 104.4, pp. 817–871.

Ayres, Ian and Peter Siegelman (1995). "Race and Gender Discrimination in Bargaining for a New Car". In: *American Economic Review* 85.3, pp. 304–321.

Bagwell, Kyle (2007). "Handbook of Industrial Organiziation". In: ed. by M. Armstrong and R. Porter. Vol. 3. Elsevier. Chap. 28.

Becker, Gary S. (1957). *The Economics of Discrimination*. University of Chicago Press.

Bessendorf, Anna (2015). *From Cradle to Cane: The Cost of Being a Female Consumer*. Tech. rep. NYC Consumer Affairs.

Blau, Francine D. and Lawrence M. Kahn (Sept. 2017). "The Gender Wage Gap: Extent, Trends, and Explanations". In: *Journal of Economic Literature* 55.3, pp. 789–865. DOI: 10.1257/jel.20160995. URL: https://www.aeaweb.org/articles?id=10.1257/jel.20160995.

Bronnenberg, Bart et al. (2015). "Do Pharmacists Buy Bayer? Informed Shoppers and the Brand Premium". In: *The Quarterly Journal of Economics* 130.4, pp. 1669–1726.

Busse, Meghan R., Ayelet Israeli, and Florian Zettelmeyer (2017). "Repairing the Damage: The Effect of Gender and Price Knowledge on Auto-Repair Price Quotes". In: *Journal of Marketing Research* 54.1, pp. 75–95.

Consumer Reports (2010). *Men Win the Battle of the Sexes*. (Visited on 2010).

DellaVigna, Stefano and Matthew Gentzkow (2020). "Uniform Pricing in U.S. Retail Chains". In: *The Quarterly Journal of Economics*, pp. 2011–2084.

Electronic copy available at: https://ssrn.com/abstract=3882214

Duesterhaus, Megan et al. (2011). "The Cost of Doing Femininity: Gendered Disparities in Pricing of Personal Care Products and Services". In: *Gender Issues* 28, pp. 175–191.

Duffin, Karen (2019). *The Problem with the Pink Tax*. (Visited on 12/06/2019).

Goldberg, Pinelopi Koujianou (1996). "Dealer Price Discrimination in New Car Purchases: Evidence from the Consumer Expenditure Survey". In: *Journal of Political Economy* 104.3, pp. 622–654.

Goldsmith-Pinkham, Paul and Kelly Shue (2020). "The Gender Gap in Housing Returns". In: *WP*.

Gonzalez Guittar, Stephanie et al. (2021). "Beyond the Pink Tax: Gender-Based Pricing and Differentiation of Personal Care Products". In: *Gender Issues*.

Hitsch, Gunter, Ali Hortaçsu, and Xiliang Lin (2021). "Prices and Promotions in U.S. Retail Markets". In: *Working Paper*.

Kahneman, Daniel, Jack L. Knetsch, and Richard Thaler (1986). "Fairness as a Constraint on Profit Seeking: Entitlements in the Market". In: *The American Economic Review* 76.4, pp. 728–741. ISSN: 00028282. URL: http://www.jstor.org/stable/1806070.

Manson, Steven et al. (2017). *IPUMS National Historical Geographic Information System: Version 12.0 [Database]*. Minneapolis: University of Minnesota.

Shapiro, Bradley, Gunter Hitsch, and Anna Tuchman (Forthcoming). "TV Advertising Effectiveness and Profitability: Generalizable Results from 288 Brands". In: *Econometrica*.

Shapiro, Carl (1982). "Product Differentiation and Imperfect Information: Policy Perspectives". In: *FTC Working Paper* 70.

Electronic copy available at: https://ssrn.com/abstract=3882214