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

Sarah Moshary
University of Chicago – Booth

Anna Tuchman
Northwestern University – Kellogg

Natasha Bhatia
Cornerstone Research

October 29, 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 provides systematic evidence on price disparities for personal care products targeted at different genders using a national dataset of grocery, convenience, drugstore, and mass merchandiser sales, in combination with novel sources on product gender targeting. We do not find evidence of a systematic pink tax: women's products are more expensive in some categories (e.g., deodorant) but less expensive in others (e.g., razors). Further, in an apples-to-apples comparison of women's and men's products with the same active and inactive ingredients, the women's variant is less expensive in five out of six categories. Our results call into question the need for and efficacy of recently proposed and enacted federal and state legislation mandating price parity across gendered products in posted price markets.

1 Introduction

The "pink tax" refers to an alleged empirical regularity that goods marketed toward women are more expensive than their counterparts marketed toward men. Gender-based

*We would like to thank Sara Drango, Malika Korganbekova, and Jason Zhao for excellent research assistance.
Researchers' 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 researchers 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.
Researchers' 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 researchers 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: sarah.moshary@chicagobooth.edu; anna.tuchman@kellogg.northwestern.edu; nbhatia@cornerstone.com

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 are keen to address these perceived inequalities through legislation. For example, in 2019-2020, the NY State Assembly and State Senate passed bill S2679 that bans pricing on the basis of gender. Since 2015, Congresswoman Jackie Speier has introduced the Pink Tax Repeal Act four times, with the goal of implementing a similar ban nationwide.

Unfortunately, there is a dearth of systematic evidence on the the pink tax to guide legislative action. Gender-based pricing likely operates differently in consumer packaged goods compared to other settings studied in the literature, which typically involve price negotiation: 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 transactions (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 secure 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, where 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. To prevent arbitrage, firms must therefore differentiate the products targeted at different genders in such a way that consumers self-select into the product designed for their group. As an example, a soap manufacturer might sell two versions of an otherwise identical soap, a low-priced blue bar and a high-priced pink bar. Second degree price discrimination of this sort can be profitable for firms if men and women have different demand for soap. Gender-based price discrimination in CPG would therefore manifest as differences in shelf price across products that are targeted at different genders, rather than differences in prices charged to men and women for the same exact product.

It is difficult to measure the pink tax precisely because gender segmentation in CPG necessitates product differentiation. If men's and women's products differ in appearance and content, then any price differences could reflect differences in markups or costs. To this end, proposed and adopted legislation bans price differences for "substantially similar" products, but does not provide clear criteria for evaluating similarity. The legislation does rule out color as a meaningful differentiator, with an exception in cases where color generates cost disparities. While the legislation leaves room for interpretation, these provisions indicate

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

that regulators seek to crack down on differential markups, while permitting differential pricing that is driven by costs.

The goal of this paper is to assess the prevalence of gender-based pricing for CPG goods. We begin by evaluating existing evidence on the pink tax from Bessendorf (2015), a New York City Department of Consumer Affairs (NYC DCA) report frequently referenced in state and national pink tax legislation. Bessendorf (2015) hand-collected prices for 122 products at three NYC drugstores and found that women's products were more expensive in six of seven personal care categories. When we compare prices of the same products but extend the analysis to include other retail formats and retail outlets across the country, we too find that average prices are higher for women's products in those same six categories. However, we hesitate to generalize these findings to the rest of the personal care market for two reasons. First, the products considered in the report account for less than 6% of category sales and were not selected at random. Second, while the sample was constructed by manually pairing men's and women's products, we find that most pairs in the sample differ in their ingredients.

We then construct our own estimates of gender price differences for a wide array of personal care products at thousands of retail outlets across the United States from 2015-2018 using Nielsen RMS data. In our first set of estimates, we do not condition on product attributes. Thus, we term this difference the *pink gap* because it could reflect either differences in markups and/or costs between the products targeted at men and women. Of the nine categories in this comparison, unit prices for women's products are higher than those for men's products in only four: bar soap, body wash, deodorant, and razor blades. In three of the other categories, the unit prices for men's products are higher than the unit prices for women's products, and the remaining two categories do not have significant differences in unit prices. In other words, while average prices for men's and women's products differ, the pink gap is often negative.

Price differences shrink even more when we refine the comparison to substantially similar products, which is how current and proposed legislation conceptualize the pink tax. Following the Pink Tax Repeal Act, we operationalize substantial similarity as products made by the same manufacturer that contain the same leading ingredients. Constructing such an apples-to-apples comparison is important to rule out the possibility that price differences stem from differences in quality and therefore costs across gendered products. For example, women might face the same prices as men but their products could be lower quality and cost less to produce. In this case, equal shelf-prices could mask a larger markup on women's goods. To the contrary, when we consider within-manufacturer comparisons of products with similar ingredients, price differences decrease further, and women's products are more expensive than men's in only one of the six categories where ingredient information is observed. Furthermore, pooling the apples-to-apples estimates across categories, unit prices for women's products are 5% cheaper than for men's products. Our findings imply that the Pink Tax Repeal Act is unlikely to meaningfully change average prices in personal care; we show that men and women already face similar prices for similar products.

The paper proceeds as follows: Section 2 describes current pink tax legislation and Section 3 details the data. Section 4 provides a replication and evaluation of existing evidence on the pink tax from the NYC DCA. Section 5 describes our preferred estimates of the pink gap and pink tax. Section 6 discusses policy implications.

2 Current Legislation

The Pink Tax Repeal Act is the principle federal legislation aimed at combating price discrimination against women in CPG. The act was first introduced in 2015 by Congresswoman Jackie Speier, who succinctly describes the act as:

> "prohibit[ing] the sale of substantially similar goods or services that are priced differently based on gender, allow[ing] the Federal Trade Commission to enforce violations, and ensur[ing] that State Attorneys General have the authority to take civil action on behalf of consumers wronged by discriminatory practices."

In practice, the Pink Tax Repeal Act defines similar products as those produced by the same manufacturer and that have no substantial differences in (a) the materials used in the products, (b) the intended use of the products, or (c) the "functioning and features" of the products. The bill specifies that differences in color *do not* qualify as substantial. The legislation has 48 current signatories and is endorsed by Consumer Reports, the Consumer Federation of America, and the National Women's Law Center.[2] To motivate the bill, Congresswoman Speier cites a report by the NYC Department of Consumer Affairs (Bessendorf, 2015) that found substantial price differences between the prices of men's and women's products. The NYC DCA study also features in the Joint Economic Committee 2016 report on the pink tax.[3] We replicate and extend estimates from this study in Section 4. Other studies of the pink tax include contemporaneous work by Gonzalez Guittar et al. (2021), which finds mixed results in their study of scraped price data from four online retailers, and a 2018 study by the Government Accountability Office, which finds that average prices are higher for women's (men's) products in five (two) of ten personal care categories.[4]

Both California and New York have passed separate legislation aimed at eradicating the pink tax. The 1996 California Gender Tax Repeal Act bans gender-based pricing of consumer services, such as haircuts and dry cleaning. As Part of the FY 2021 Budget Bill,

[2] `https://speier.house.gov/press-releases?id=C2F060D1-0D84-4824-B9E5-40F879F22CFA`

[3] Report available at: `https://www.jec.senate.gov/public/_cache/files/8a42df04-8b6d-4949-b20b-6f40a326db9e/the-pink-tax---how-gender-based-pricing-hurts-women-s-buying-power.pdf`.

[4] The report can be found at: `https://www.gao.gov/products/gao-18-500`.

Governor Andrew Cuomo of New York included a provision expressly banning the pink tax. Similar to the Pink Tax Repeal Act, the bill defines the pink tax as gender-based pricing for substantially similar products. The New York ban went into effect September 30, 2020. The bill describes exemptions when differences in prices reflect differences in cost.[5]

3 Data

Retail Prices

We use Nielsen Retail Scanner data from 2015 to 2018 to document price differences between personal care products targeted at men and women. We examine nine categories: bar soap, body wash, deodorant, hair coloring,[6] razor blades, disposable and non-disposable razors, shampoo, 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. A feature of the data is that it reflects price promotions, whereas other studies of the pink tax, such as Bessendorf (2015), focus on the full sticker price. The data also includes product characteristics, such as brand name 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 (price charged). First, we assume that a product with zero sales in a particular week was offered at its regular (non-discounted) 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

We extract information on product-level gender-targeting from the following sources:

1. **Nielsen Brand and Product Module Descriptions:** We search for gendered words such as "his" and "hers" in Nielsen's brand description for each UPC and product module name.[7]

[5] `https://www.governor.ny.gov/news/governor-cuomo-reminds-new-yorkers-pink-tax-ban-goes-effect-today`

[6] 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.

[7] To classify men's UPCs, we searched for the following words: his, 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: her, lady, girl, ldy, women, femme, ladies, lady's, and wmn.

2. **Label Insight:** We collect data on gender targeting from Label Insight, a market research firm that records marketing claims for CPG brands. The database also includes product pictures.[8]

3. **Walgreens Website:** We scrape gender categorizations from `Walgreens.com`, the website of the large American drugstore chain. Web Appendix B.1 displays screenshots of gender categorization and filters on the Walgreens webpage. Scraping was performed in Summer 2020.

4. **Differential purchasing by all-male and all-female households in the Nielsen consumer panel dataset from 2006 to 2018.** We identify products whose consumer base is significantly skewed towards one gender using data from the Nielsen Consumer Panel on the purchases of single-gender households. These households account for more than 25% of households in the panel.[9] For each UPC, we define the female (male) share as the percent of 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 via a binomial test where the null hypothesis is that the female (male) share is equal to 71% (29%).[10] If we do not reject the null, the product is left uncategorized. Skew in purchasing could indicate an explicit gender cue (e.g., a label or picture), or simply an attribute with a gender-specific appeal. Empirically, these two sets turn out to be similar but not identical.[11]

5. **Hand-coding Label Insight product images.** We hired undergraduate research assistants at the University of Chicago and Northwestern University to categorize product images from Label Insight. Products are categorized as male, female, unisex, or unknown. Web Appendix B.2 describes the recruiting and labeling process in detail.

We combine these sources to construct a single gender variable. 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.[12] 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.

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

[9] Female-only households are more common – they represent 71% of single-gendered households.

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

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

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

Table 1: Gender Targeting across Personal Care Categories

Nielsen Product Module	% Qty Gendered All	% Qty Gendered for Women	Total Qty (MM)	% UPCs Gendered All	% UPCs Gendered for Women	Count of UPCs
Soap - Bar	71.7%	57.3%	523	19.1%	54.7%	3,710
Soap - Liquid	46.3%	94.2%	508	13.7%	90.6%	3,100
Soap - Specialty	78.5%	63.0%	820	26.7%	58.8%	6,889
Deodorants - Personal	99.0%	49.7%	1,059	76.6%	46.1%	2,958
Hair Coloring	100.0%	88.9%	312	100.0%	95.8%	2,534
Hand & Body Lotions	73.4%	95.4%	472	18.7%	89.2%	7,593
Razor Blades	86.5%	32.9%	96	51.4%	39.3%	519
Razors Disposable	68.2%	51.2%	289	36.5%	49.9%	978
Razors Non-Disposable	88.5%	45.8%	68	48.8%	38.2%	451
Creme Rinses & Conditioners	81.1%	99.5%	551	26.4%	93.6%	5,916
Shampoo						
–Aerosol/ Liquid/ Lotion/ Powder	71.1%	74.3%	895	26.8%	68.3%	7,746
–Bars/ Concentrates/ And Creams	65.9%	89.1%	9	21.4%	89.3%	131
–Combinations	49.8%	99.9%	20	18.1%	83.0%	519
Shaving Cream	100.0%	25.2%	271	100.0%	21.9%	942

Notes: This table describes the share of products available at Nielsen RMS stores between 2015-2018 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.

Table 1 shows the pervasiveness of gender targeting across product modules.[13] 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 personal care products in the Nielsen data, although there is considerable variation 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. One reason for the substantial share of products with no gender label is that it is challenging to label niche products. Market shares reveal that gendered products account for 80% of volume sales across all categories.[14]

[13] 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.

[14] 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 B.4). 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.

Ingredients

We use Syndigo Product Label Data on product ingredients. For each UPC, the data includes the names and amounts of any active ingredients as well as the names of inactive ingredients and their relative prevalence in the product. The data covers 7,020 of the 25,982 personal care products in our sample of gendered products from the Nielsen data.

4 NYC Replication

In this section, we revisit evidence from Bessendorf (2015), a NYC Department of Consumer Affairs study that reports a 13% pink tax in personal care. We focus on this report because it is cited as motivation both for proposed federal legislation and existing state regulation on the pink tax. We begin by replicating the results of the report using the original data collected by the NYC DCA for the study. This data was collected for 122 UPCs sold in NYC drugstores from July-October 2015. Next, in order to understand whether the 13% price difference is peculiar to New York City or represents a broader phenomenon, we extend the scope of the analysis by examining the prices charged for these same products by a large sample of supermarkets, mass merchandisers, convenience stores, and drugstores across the US. We then provide evidence on the comparability of the men's and women's products studied in the report.

Table 2 reports estimates of price disparities calculated following Bessendorf (2015). The report measures the so-called "pink tax" by pairing men's and women's products, calculating the within-pair price difference, averaging price differences across pairs within a category, and then scaling by the average price for men's products in the category. In cases where the men's and women's products are different sizes, the report re-scales prices using the ratio of sizes (e.g., by multiplying the price per oz of the men's product by the size of the women's product).[15] It arrives at a 13% pink tax via a simple average across categories. Column (6) replicates Bessendorf (2015)'s estimates using the original data collected by the NYC DCA.[16] Based on the NYC DCA data, women's products are more expensive in five out of six personal care categories. Our aim is to understand whether and to what extent these price differences extend to other stores, retail formats, and geographies. Using the Nielsen RMS data, we estimate price differences for the set of products (UPCs) considered in the report for three samples: drugstores in New York City, all drugstores, and all retailers. Our analysis

[15] The report does not rescale prices for body wash. Because our aim is to replicate their methodology, estimates in Table 2 do not rescale in this category either.

[16] We are able to replicate all values in Table 5 in Bessendorf (2015) except the average price of razors targeted to women. Table 5 in Bessendorf (2015) reports an average price of \$8.90 for women's razors, while we find an average price of \$8.73. This difference in price leads them to report a price gap of 11% while we compute a price gap of 9%. We believe the discrepancy is most likely due to a typo in the product-level price data reported in Bessendorf (2015)'s appendix or a mistake in computing the averages reported in their summary table.

Table 2: Replication and Extension of NYC DCA Report Pink Tax Estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
			Estimate	Men's Price	Pink	NYC Report Estimates	
Category	Channels	Geographies	($)	($)	Gap	Reported	Nielsen UPCs
Body Wash	Drugstores	NYC Only	0.45***	5.73	7.9%	5.5%	5.5%
	Drugstores	National	0.54***	4.85	11.1%		
	All	National	0.74***	4.53	16.4%		
Deodorant	Drugstores	NYC Only	0.06***	5.15	1.1%	3.3%	4.0%
	Drugstores	National	0.31***	4.27	7.2%		
	All	National	0.40***	3.90	10.1%		
Hair Care	Drugstores	NYC Only	2.35***	7.88	29.9%	47.7%	29.7%
	Drugstores	National	0.80***	6.38	12.6%		
	All	National	0.22***	5.09	4.3%		
Razor	Drugstores	NYC Only	0.78***	10.60	7.4%	9.3%	15.2%
	Drugstores	National	1.53***	8.51	18.0%		
	All	National	1.18***	8.54	13.9%		
Razor Cartridges	Drugstores	NYC Only	2.59***	15.34	16.9%	10.9%	15.4%
	Drugstores	National	2.11***	14.06	15.0%		
	All	National	2.21***	14.09	15.7%		
Shaving Cream	Drugstores	NYC Only	-0.48***	4.09	-11.7%	-4.1%	-13.0%
	Drugstores	National	-0.38***	3.67	-10.5%		
	All	National	-0.35***	3.46	-10.0%		

Notes: The pink tax is measured as the ratio of the estimated price difference (column (3)) to the average price of a men's product in the same category (column (4)) multiplied by 100. Columns (6) and (7) present estimates of the pink tax using the NYC DCA data, where column (7) subsets to the products that can be matched to the Nielsen data. These prices of these matched products in the Nielsen data comprise the sample in columns (3)-(5).

excludes 33 of the products in the NYC DCA sample (27 of them are private label products) because we cannot match them to a product observed in the Nielsen data.[17] However, we do not believe this substantively affects our estimates of price differences; column (7) shows that this subset of UPCs produces similar estimates of the pink gap in the NYC DCA data. Column (3) reports average price difference in dollars for different samples, and column (5) reports the implied pink tax. The estimates echo Bessendorf (2015) in that five of the six categories feature a price premium for women's products.

We next consider the generalizability of these estimated price gaps beyond the products studied in Bessendorf (2015). The question of extrapolation is important because the products in the sample comprise less than 6% of category sales and were not selected at random.[18] Rather, the sample was constructed by manual identification of men's and women's products that were perceived to be comparable. Correctly constructing an apples-to-apples comparison is important to ensure that estimated price differences do not reflect differences in marginal cost and also to evaluate proposed legislation, which mandates price parity only in instances where men's and women's products are substantially similar. The report does not provide its criteria for comparability, and perusal of product pairs included in the report reveals salient differences: as an example, in two of eight shampoo comparisons, the price of a single 2-in-1 men's product is compared to the combined price of a women's shampoo and a women's conditioner, producing price gaps over 100%. To provide systematic evidence on the similarity of product pairs, we leverage data from Syndigo on product ingredients. Table 3 reports the number of pairs in each category with matching ingredients. The criteria for matching ingredients becomes more stringent from left to right in the table; column (3) reports the number of pairs with the same active ingredient (which is only relevant in certain categories), column (4) reports the number with the same active and first inactive ingredients, etc.[19] Only one-third of product pairs comprise the same leading ingredients. We note that the challenge of identifying similar products is compounded by the challenge of identifying gender targeting. The NYC DCA report includes comparisons between explicitly labeled men's products and unisex products in cases where no women's product could be identified. These issues of comparability in Bessendorf (2015) hamper interpretation of the price difference estimates in Table 2 as a pink tax. It is unclear whether the estimates reflect differences in the attributes of men's and women's products or differences in the mapping from attributes to prices for men's and women's products (i.e., markups) and, more broadly, whether the 122 products considered are representative of personal care. In the next section, we describe our preferred methodology for calculating price disparities in personal care

[17] Because the comparison studies only within-pair prices, we lose a further 9 products that do not have a match in our sample.

[18] Appendix table A2 shows the market share of products in the NYC DCA sample by category.

[19] The FDA mandates that active ingredients are reported first, then inactive ingredients in descending order of predominance, and then any order is permitted for ingredients that comprise less than 1% of the product. [https://www.fda.gov/cosmetics/cosmetics-labeling-regulations/cosmetics-labeling-guide#clgl]

Table 3: Similarity of Product Ingredients for NYC DCA Report Product Pairs

Product Category	N Pairs	N Pairs w/ Active	N Pairs Matching Up To Active	Inactive 1	Inactive 2	Inactive 3	Inactive 4	Inactive 5
Body Wash	9	0	-	9	7	7	5	2
Deodorant	9	9	9	9	9	6	6	6
Hair Care	6	2	1	5	3	2	2	2
Shaving Cream	6	0	-	6	4	4	1	0
Total	30	11	91%	97%	77%	63%	47%	33%

Notes: Column (1) reports the number of product pairs that we could identify in the Nielsen data and column (2) reports the number of pairs that have an active ingredient. The remaining columns report the number of pairs that match up to and including that ingredient. For example, the last column reports the number of pairs that match on active ingredient and the first five inactive ingredients.

that leverages a national dataset of prices and product ingredients, as well as a data-driven approach to identifying gender targeting.

5 Measuring Price Disparities

We estimate the unconditional gender price gap as the difference in the average price of products targeted at men and women across retail outlets from 2015-2018. This specification includes all gendered products and so reflects both differences in the attributes of products targeted at men and women and differences in the mapping between attributes and prices for men's and women's products. To the extent that firms select the attributes of gendered products to segment consumers, these price differences can therefore incorporate both second and third degree price discrimination.

Our main specification models 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 (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 pink gap. We estimate equation (1) separately for each personal care product module.[20]

In contrast with Bessendorf (2015), we find that women face lower average shelf prices in a majority of personal care product modules. Specifically, we estimate equation (1) with shelf price per product as the dependent variable, where observations are weighted by the number of weeks a product is 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,[21] and deodorant products targeted at women are more expensive than their counterparts targeted toward men, but hair coloring, razors and razor blades, shampoos, and shaving creams targeted at women are less expensive.

A key difference between these estimates of the pink gap and those presented in Bessendorf (2015) is the treatment of product size. If women's products tend to be smaller than men's products, then lower product shelf prices might obfuscate higher per-unit prices. To address this concern, we repeat this exercise using unit shelf price as the dependent variable (i.e. price per ounce or count, depending on the category). Results are reported in column (2): on a per unit basis, women's (men's) products are more expensive in four (three) of nine product categories. The remaining two categories do not have statistically significant differences in unit prices. While the split of categories seems fairly even between those where women fare better/worse, we note that the magnitudes of price differences are larger in the categories where women's products are more expensive. For example, the average unit price of women's bar soap is 110% more expensive than the average unit price of men's bar soap, while the average unit price of women's shaving cream is 17% less expensive than the average unit price of men's shaving cream.

We next consider the difference in the price paid for men's and women's products. This measure gives a sense for how differences in offered prices translate to differences in expenditures and whether price differences create an economically significant burden for women. This analysis is novel in the pink tax literature. 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 unit price paid for women's and men's products, which we obtain by estimating equation (1) using quantity sold as regression weights. We interpret these quantity-weighted regressions as evidence on the prices that men and women pay for personal care products because we find that women seldom purchase products targeted at men and vice versa; as an example, 78% of women and 81% of men that participate in the Nielsen panel only purchase

[20] We restrict our analysis to product modules that are not dominated by a single gender and modules with sufficient sales volume. Specifically, we focus on modules in which neither gender accounts for more than 90% of gendered volume sales and that have at least 10 million units sold.

[21] Called "(Soap - Specialty)" in the Nielsen RMS data.

Table 4: Price Gap by Category, 2015-2018

Module	(1) Product Shelf Price	(2) Unit Shelf Price	(3) Unit Price Paid	(4) Unit Shelf Price	(5) Unit Shelf Price
Bar Soap	0.37***	0.25***	0.22***	0.28***	0.03***
	(0.04)	(0.00)	(0.00)	(0.01)	(0.00)
	4.31	0.23	0.21	0.23	0.23
	8.7%	110.3%	109.0%	122.6%	14.0%
Body Wash	0.73***	0.18**	0.14**	0.11***	-0.02***
	(0.06)	(0.03)	(0.02)	(0.01)	(0.00)
	4.57	0.28	0.26	0.28	0.28
	16.0%	64.3%	53.1%	40.5%	-6.5%
Deodorant	0.29***	0.50***	0.44***	0.50***	-0.07***
	(0.02)	(0.02)	(0.02)	(0.01)	(0.00)
	4.62	1.51	1.38	1.46	1.46
	6.3%	33.1%	31.6%	34.2%	-4.5%
Hair Coloring	-0.81***	-0.52**	-0.71**	-0.69**	0.21
	(0.03)	(0.13)	(0.17)	(0.13)	(0.28)
	8.78	8.48	8.07	8.78	8.78
	-9.2%	-6.1%	-8.8%	-7.9%	2.4%
Razor Blades	-4.14***	0.70***	0.59***		
	(0.31)	(0.04)	(0.05)		
	22.02	3.67	3.64		
	-18.8%	18.9%	16.2%		
Razors Disposable	-0.62***	-0.18*	-0.17**		
	(0.04)	(0.06)	(0.03)		
	7.72	2.26	2.05		
	-8.1%	-7.9%	-8.1%		
Razors Non-Disposable	-1.05**	-1.05**	-0.95**		
	(0.21)	(0.21)	(0.25)		
	11.69	11.69	11.18		
	-9.0%	-9.0%	-8.5%		
Shampoo	-1.09**	0.00	0.02	-0.01	-0.06**
	(0.20)	(0.01)	(0.01)	(0.01)	(0.01)
	6.33	0.49	0.43	0.48	0.48
	-17.2%	0.6%	4.5%	-1.6%	-12.1%
Shaving Cream	-0.42***	-0.10***	-0.07***	-0.08***	-0.04***
	(0.04)	(0.01)	(0.01)	(0.01)	(0.01)
	3.60	0.59	0.51	0.54	0.54
	-11.6%	-17.4%	-13.6%	-14.2%	-8.0%
Data	All	All	All	Syndigo	Syndigo
Ingredients FE	N	N	N	N	Y
Manufacturer FE	N	N	N	N	Y

Notes: The sample in columns (1)-(3) comprises the full set of products. Column (4)-(6) subset to products with observed ingredient and manufacturer information. 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 to row three. Regressions are estimated separately by product module and include store and year fixed effects. Columns (4)-(5) exclude razors because their ingredients are not reported. *** $p < 0.01$, ** $p < 0.05$, * $p < .01$.

deodorants targeted at their own gender.[22] The estimates imply that women pay higher prices for bar soap, body wash, deodorant, and razor blades on a per unit basis, while men pay higher prices for hair coloring, disposable and non-disposable razors, and shaving cream. In seven of the nine product categories, the magnitude of the differences in price paid is smaller than the magnitude of the differences in price charged.

We now return to the question of comparability between men's and women's products. The estimates in columns (1)-(3) show that women's products are less expensive than men's products in several categories, but they do not rule out differential markups for women's products in those same categories. In other words, the results so far do not preclude the possibility that women's products are lower quality, but are sold at higher markups. Recognizing that the features of men's and women's products may differ, proposed legislation mandates price parity specifically in instances where men's and women's products are substantially similar and made by the same manufacturer; it is price differences for these comparable pairs that legislation defines as the pink tax. The legislation provides only loose guidance on how to determine substantial similarity, so we assess similarity on the basis of product ingredients. The law does explicitly restrict to within-store price comparisons, limiting its scope. Table 5 describes the extent of overlap in manufacturer-formulations across genders within store.[23] As an example, Table 5 shows that the average retail outlet carries 32 unique deodorant formulations targeted to women and 34 targeted to men. Approximately 28% of formulations targeted to women have a comparable formulation targeted to men within the same store and vice versa. There is substantial variation across categories: hair coloring formulations have almost no overlap across genders, and deodorants have the most overlap. Turning to unique UPCs, in the average store less than half of gendered UPCs have a comparable product targeted at the other gender. These patterns indicate that legislation like the Pink Tax Repeal Act would not apply to most personal care products.

We estimate the pink tax for substantially similar products by incorporating data on manufacturer and product ingredients into equation (1). This exercise subsets to the sample of Nielsen products that merge to the Syndigo data on product ingredients. This merged sample is similar to the full set of Nielsen products in terms of the unconditional pink gap, as shown in column (4) in Table 4, which replicates the regression in column (3) for products with ingredient information. Column (5) displays price gaps when we add fixed effects for

[22] These statistics are based on the purchase of products where gender is defined using all data sources, but the same patterns hold if we rely only on explicit gender targeting that excludes the panelist data.

[23] This exercise is conducted for the subset of products with Syndigo ingredient data. A product formulation is defined as the combination of manufacturer, active ingredient and the first five inactive ingredients, where the order of ingredients matters. In using this subset of the ingredients, our goal is to identify products that are similar in overall formulation without requiring otherwise similar products to be identical on more cosmetic features. We conduct the analysis on 2018 data to ensure similar formulations were available at the same point in time, and we exclude convenience stores from this analysis because these stores carry very limited assortments.

Table 5: Overlap in Manufacturer-Ingredients Across Genders, 2018

Module	Gender	(1) N Formula	(2) % Formula	(3) N UPCs	(4) % UPCs	(5) Unit Sales	(6) % Sales
Bar Soap	men	9	14.5%	19	14.9%	962	11.6%
	women	12	12.0%	24	13.9%	1,239	15.0%
Body Wash	men	19	16.3%	32	16.5%	1,278	23.3%
	women	38	8.5%	56	15.9%	1,973	25.6%
Deodorants	men	34	28.1%	85	30.8%	3,213	25.5%
	women	32	28.5%	80	47.8%	2,817	44.1%
Hair Coloring	men	7	0.1%	15	0.1%	243	0.0%
	women	32	0.0%	142	0.0%	1,514	0.0%
Shampoo	men	16	25.6%	29	30.3%	791	31.7%
	women	49	7.5%	73	18.0%	1,884	23.0%
Shaving Cream	men	12	12.5%	21	13.5%	971	13.1%
	women	6	28.3%	9	37.6%	290	43.5%

Notes: Columns (1), (3), and (5) report the number of unique formulations, number of UPCs, and the unit sales for the average store in 2018. We define a formulation as the combination of manufacturer, active ingredient, and top five inactive ingredients. Column (2) reports the fraction of formulations targeted to one gender for which there is a comparable formulation targeted to the other gender. Column (4) reports the fraction of UPCs targeted to one gender for which there is a comparable formulation targeted to the other gender. Column (6) reports the fraction of unit sales for one gender's products for which there is a comparable formulation targeted to the other gender. The analysis is conducted on the subset of products with ingredients information in the Syndigo data and convenience stores are dropped because they have very small assortments.

the manufacturer, active ingredient, and the first five inactive ingredients. The gender price gap shrinks towards zero in all but one product category. Controlling for manufacturer and ingredients, women face lower prices in four of the six product categories; men face higher prices for body wash, deodorant, shampoo, and shaving cream, while women face higher prices for bar soap.

To compute the overall price difference for personal care products, we re-estimate equation (1) stacking the data across categories, weighing categories by assortment size, and using log price per unit as the dependent variable. Unconditional on attributes, unit prices for women's products are 18% more expensive than unit prices for men's products ($SE = 0.001$, $p < 0.01$), but conditional on manufacturer and ingredients, unit prices for women's products are on average 5% cheaper ($SE = 0.01$, $p < 0.05$).

Taken together, the estimates in Table 4 do not support the hypothesis that women systematically face and pay higher prices in personal care. Rather, they tell a more complex story: there are economically and statistically significant price differences across genders, but the direction of these differences varies across product categories. Furthermore, price differences shrink when we look within substantially similar products, which is the focus of current legislation. 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, retailers, and geographies.

6 Discussion

Gender inequalities in the labor market have spawned a substantial literature in economics and a spate of federal and state regulations.[24] In this paper, we show that recent concern that gender price discrimination extends to personal care products is unfounded. We leverage a national dataset of prices and sales of grocery, convenience, drug, and mass merchandise retail outlets coupled with detailed data on product gender-targeting and ingredients to shed light on the pink tax. We begin by evaluating existing evidence on the pink tax from Bessendorf (2015), which compares 122 seemingly similar men's and women's products sold in NYC drugstores. Focusing on the same set of products considered in the report, we find similar price differences using a national panel of retail chains. However, we find that the men's and women's products in the study differ in their ingredients, hampering interpretation of the reported price differences as a pink tax. Further, the products in the report represent only a small share of category sales and exclude many big national brands.

We construct our own measure of price differences to provide systematic evidence on gender-based pricing in CPG. First, we estimate the difference in the average price of men's and women's products within-store and time period. We term this comparison the "pink gap" rather than the "pink tax" because it may confound differences in markups with differences

[24]E.g., the Equal Pay Act of 1963.

in marginal costs, stemming, for instance, from differences in product attributes. We find that the pink gap is often negative; men's products command higher per-product prices in six of nine categories that we study and higher unit prices in three of nine categories. We then estimate the pink tax via an apples-to-apples comparison of products manufactured by the same firm and comprising the same leading ingredients. Controlling for attributes shrinks price differences in all but one category. Further, men's products are more expensive in four of six categories when we control for ingredients. Taken together, our findings do not support the existence of a systematic price premium for women's products, but our results do reveal that gender segmentation in personal care is pervasive and operates through product differentiation. A back-of-the-envelope calculation implies that the average household would save less than 1% by switching to substantially similar products targeted to a different gender. The potential savings are much larger – on the order of 20% – if a household were willing instead to substitute to products with different gender-targeting and different formulations.[25] However, a revealed preference argument suggests that such switching would lower consumer welfare.

The implications of our findings for current and proposed legislation are several. First, our finding that women's personal care products are not systematically more expensive calls into question the role of government intervention to reduce the pink tax. We acknowledge that our findings speak to average price differences, which may mask instances of a particular retail outlet pricing in a way precluded by the Pink Tax Repeal Act. As an example, if one store sets a 5% higher price for the men's version of a product and a neighboring store sets a 5% higher price for the women's version, we would detect no gender price gap, but the Pink Tax Repeal Act would require that both retailers change their pricing policy or alter their product assortments. However, our analysis reveals that most women's products do not have a men's analog sold in the same retail store, limiting the scope for such adjustments. Even in cases where a retailer does sell a women's variant at a higher price than its men's analogue, the Pink Tax Repeal Act might induce the retailer to drop the men's variant, de

[25] To approximate household savings, we first compute the dollar spending, average price, and total volume (measured in ounces or counts) of purchases made by each Homescan household for each product category/gender combination analyzed in Table 4. Next, for each household/category/gender, we construct the counterfactual price a household would pay if they were willing to switch to the cheaper gender within each product category. We do this by adjusting the household's price paid for the more expensive gender by the estimated price gaps reported in Table 4. When estimating savings from switching to a comparable formulation, we use the estimates in column (5), and when estimating savings when switching across formulations, we use the estimates in column (2). We then compute the household's counterfactual personal care spending by multiplying the counterfactual prices by the observed purchase volumes and summing across categories. When estimating savings from switching to a comparable formulation, we also need to account for whether a household's purchases actually have a formulaic analog that is targeted to the other gender. We do so by multiplying each household's category-level purchase volumes by the fraction of each gender's unit sales that have a comparable formulation on the shelf in the average store (column (6) of Table 5). The estimated savings from switching within formulation across gender (0.9%) are much lower than the potential savings from switching across formulations (20%) both because most purchases don't have a comparable formulation offered to the other gender in the same store, and because the price gap within a formulation is substantially smaller than the price gap unconditional on formulation.

facto increasing price dispersion by setting an infinite price.

Finally, while our findings do not support the existence of a pink tax as conceived by regulators, an economist might define the pink tax differently, perhaps as systematic differences in markups across men's and women's products. This alternative definition departs from the current and proposed legislation in that it classifies the following two cases as forms of the pink tax: (1) if men's products are cheaper to produce than their women's counterparts, e.g., due to economies of scale and (2) if markups for unique-to-men products are systematically lower than markups for unique-to-women products. We think the first scenario is unlikely because, as shown in Table 5 column (3), we do not observe systematic differences in assortment sizes that would engender differences in scale across genders in most categories. Regarding the second point, given our finding that similar men's and women's products have similar prices, we question whether differential markups for gender-unique products are relevant to the Pink Tax debate.[26] One might see differential markups for differentiated goods as unfair if they are sustained through frivolous or spurious attributes, as discussed in Shapiro (1982) and Bronnenberg et al. (2015). On the other hand, retailers and manufacturers may set markups in personal care to reflect differences in product performance that bring real value to customers, employing a commonplace pricing strategy that is perhaps no different than matinee pricing at movie theaters or early-bird specials at restaurants.

[26]We note that the debate over the tampon tax is related but distinct in that the outcry surrounds differential tax rates for tampons compared to other necessities rather than the markups for tampons.

References

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.

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 (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 J. et al. (2015). "Do Pharmacists Buy Bayer? Informed Shoppers and the Brand Premium". In: *The Quarterly Journal of Economics* 130.4, pp. 1669–1726. ISSN: 0033-5533. DOI: 10.1093/qje/qjv024. eprint: https://academic.oup.com/qje/article-pdf/130/4/1669/30637328/qjv024.pdf. URL: https://doi.org/10.1093/qje/qjv024.

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).

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*.

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

A Imputing Regular Price in the RMS Data

The Nielsen RMS data does not record the shelf price of a product (UPC) at a store in weeks when that store does not sell any units of that product. We impute missing prices using an algorithm that is motivated by the insititutional practice that retailers rarely change their regular shelf price for a product, and instead create short-term variation in prices by running temporary price promotions that discount off the regular price. Motivated by these institutional pricing practices, we use prices of the same product at the same store location in recent weeks to construct a "regular" price series, i.e., the price that would have been charged if no discounts were available that week. We operationalize this approach by setting the regular price to be equal to the maximum price observed in the current, preceding, and subsequent 4 weeks. In any weeks with an unobserved price, we then set price equal to the regular price. This is based on the intuition that zero-sales weeks are most likely to occur when the product is not on discount.

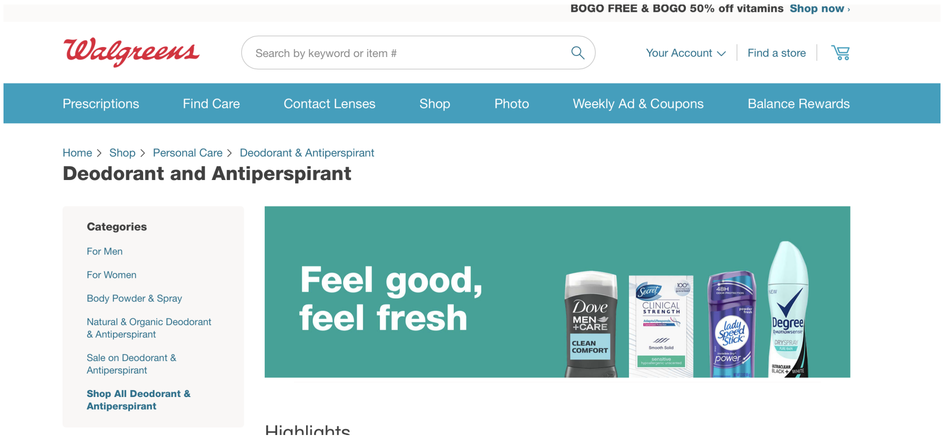
B Gender-Targeting Data Sources

B.1 Walgreens

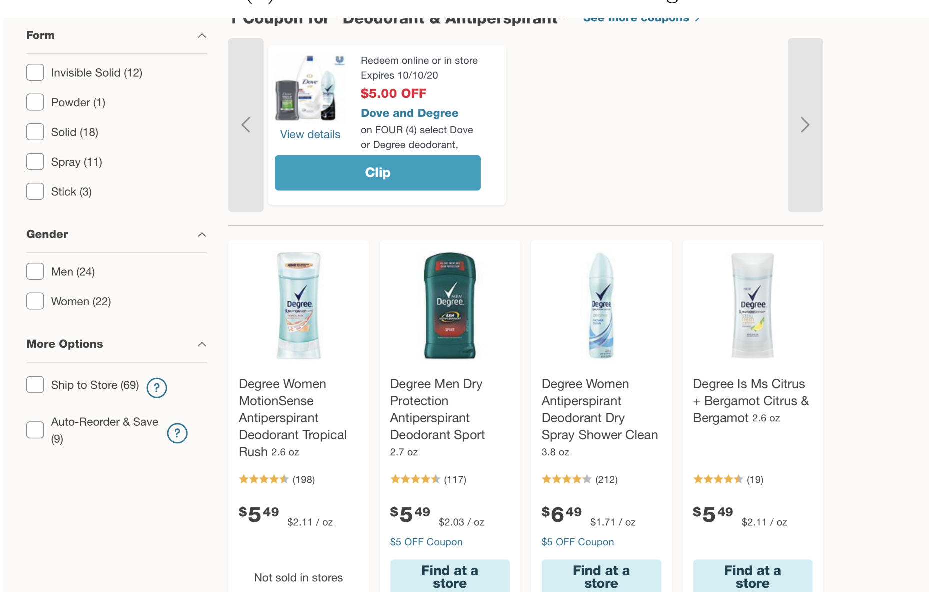
We extract gender information from the Walgreens website. The website explicitly categorizes certain product categories by gender. Figure A1 (a) presents one such example for the Deodorant & Antiperspirant category. We also collect gender information from search result page gender filters, as in Figure A1 (b).

Figure A1: Walgreens.com Gender Categorizations

(a) Primary Gender Classification



(b) Gender Filter on Search Results Page



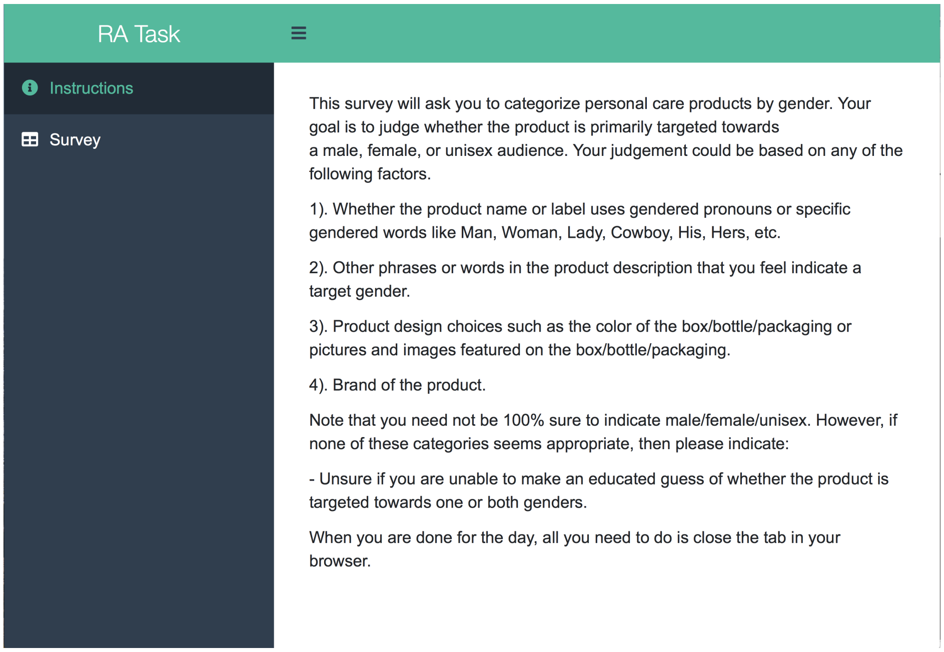
Notes: Screenshots taken on `Walgreens.com` on September 1, 2020.

B.2 Hand-Coding Product Images

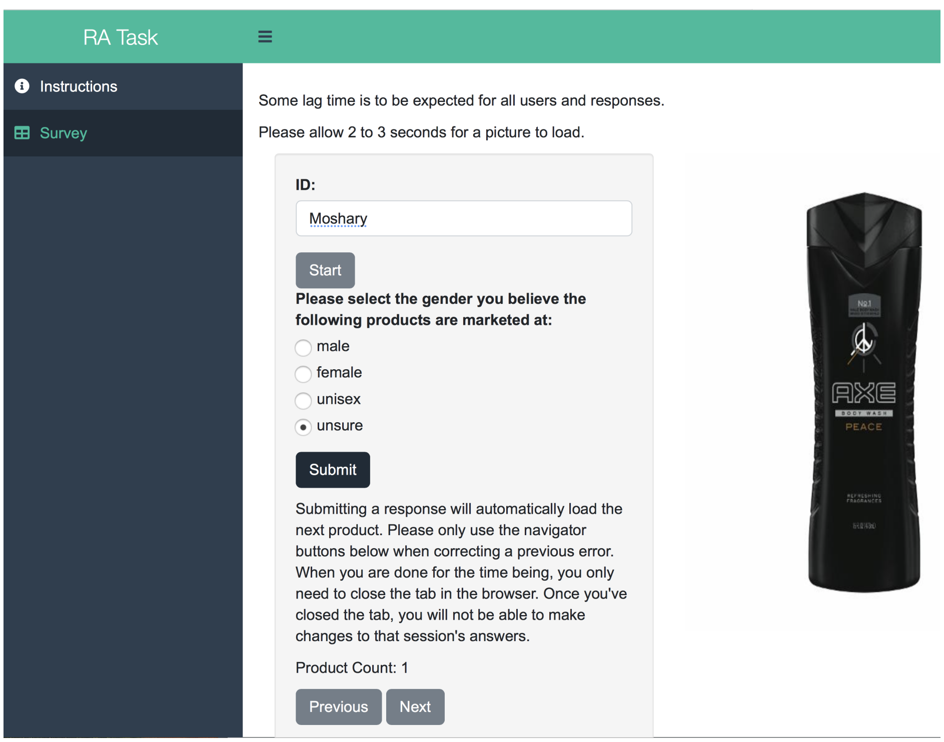
We recruited three undergraduates at the University of Chicago to assign gender labels to 1,302 personal care product images from Label Insight. The research assistants were selected based on their performance on a 25 image training dataset, where their answers were compared to our own hand coding. RAs were directed to a webapp (`https://task.shinyapps.io/classify-products/`) on September 1, 2020. Figure A2 provides snapshots of the webapp. We take the modal gender label across the three RAs; we do not record a gender label in the instances where all three RAs disagreed on their classification.

Figure A2: Webapp for Gender Classification

(a) Instructions



(b) Task



B.3 Panelist Purchases

The 2006-2018 Nielsen panelist data provides additional information on gender targeting. Intuitively, we aim to infer a product's intended gender target based on a significant skew in purchasing toward men or women. Because the data does not include the identity of the household member who purchases or consumes a product, we focus on single-gendered households for this analysis. These households comprise approximately 30% of households in the data: 14,421 all-women and 37,569 all-men households. For each product, we label products as targeted at women (men) if the share of purchases from all-women (all-men) households is significantly higher than would be expected from their preponderance in the data. Formally, we treat the number of single-gendered households that purchase an item as the number of trials in a binomial distribution, where the number of all-women (all-men) households that purchase is the count of successes. The null hypothesis in our binomial test is a one-tailed test that all-men and all-women households are equally likely to purchase the product. A product is determined to be targeted at women (men) if the null is rejected at the 5% level. This approach categorizes approximately 247,358 products (including, but not limited to, personal care). It is particularly helpful for products in early years in the sample and for products that use non-verbal cues to signal gender, such as brands like Old Spice, Secret, and Axe.

B.4 Prevalence of Private Label Products by Personal Care Category

Because Nielsen masks the UPC of private label products, we cannot identify gender targeting for these products, except through the Homescan panelist approach described in Appendix B.3. To give a sense for the importance of private label products in the personal care market, Table A1 summarizes the market share of the store brand across categories. The market shares are modest overall, with the exception of disposable razors where private label products hold a 27% market share. We acknowledge this limitation for this category.

Table A1: Market Share of Store Brand by Product Module

Nielsen Product Module	Store Brand Share
Soap - Bar	4.15%
Soap - Liquid	21.96%
Soap - Specialty	7.75%
Deodorants - Personal	0.03%
Hair Coloring	1.05%
Razor Blades	9.49%
Razors Disposable	26.75%
Razors Non-Disposable	6.98%
Creme Rinses & Conditioners	0.62%
Shampoo-Aerosol/ Liquid/ Lotion/ Powder	2.40%
Shampoo-Bars/ Concentrates/ And Creams	11.66%
Shampoo-Combinations	0.44%
Total	8.09%

Table A2: Market Share of UPCs Studied in the NYCDCA Report

	(1)	(2)
Category	UPCs Market Share	Brands Market Share
Bodywash	4.1%	32.2%
Deodorant	5.3%	35.6%
Razors	12.5%	23.2%
Shampoo	2.5%	20.7%
Shaving Cream	19.7%	48.8%
Total	5.8%	30.3%

C Additional Details on New York City Report Replication

In Section 4, we analyze the prices of products included in Bessendorf (2015) using the Nielsen RMS data. Table A2 reports the market share of these products in the Nielsen RMS data. As shown in column (1), across all categories, the share is modest, ranging from 4.1% of body wash sales to 19.7% of shaving cream sales. These figures indicate that the sample of products omits much of the personal care product landscape. This concern is amplified because the sample was not selected at random. For example, the sample omits products from some of the most popular brands because they are produced by a manufacturer uses different brand names for their men's and women's products(e.g. P&G's Secret and Old Spice brands). Column (2) reports the combined market share of brands represented in the sample, which is less than 50% for all categories.

Replicating and extending the NYC DCA analysis using the Nielsen data requires identifying the UPCs of the products in the survey, which are described on page 65 of the report. We proceed in three steps:

1. Google search for product names and descriptions. We discern the UPC from images of the back of products or from Amazon and Walmart third-party sellers. We used our best judgement in cases where product descriptions are vague.

2. For UPCs recovered in step 1, we merge to the Nielsen data using the full UPC or alternatively the UPC without the check digit. We remove any candidate matches where the Nielsen and NYC DCA report product descriptions conflict on size or brand. This left 76 matches between the NYC report and the Nielsen data.

3. For the thirteen remaining UPCs in the NYC DCA report without a match, we search for the product directly in the list of products sold in NYC drugstores in the Nielsen data.

We also emulate Bessendorf (2015) in the construction of prices using the following steps:

- In comparisons where a men's 2-in-1 shampoo and conditioner is compared to two women's products, a shampoo and a conditioner, we collapse the latter into a single observation. This requires filtering to stores and years that have both the shampoo and conditioner for a given year.

- For product pairs where the women's and men's products are different sizes, we create an "equivalent price" that is the max size within a pair multiplied by each product's unit price. Because the report does not re-scale for body wash products, we do not rescale in the body wash category.

We estimate price disparities via regressions of equivalent price on an indicator for whether the product is targeted at women. The estimates include store, year, and product-pair fixed effects.