Gender-Based Pricing in Consumer Packaged Goods: A Pink Tax?*

Sarah Moshary
University of Chicago – Booth

Anna Tuchman
Northwestern University – Kellogg

Natasha Bhatia
Cornerstone Research

January 9, 2023

Abstract

This paper investigates a controversial application of a textbook pricing practice: gender-based price segmentation, which has allegedly created a *pink tax* whereby products targeted at women are more expensive than comparable products marketed toward men. Our results shed light on the form and magnitude of gender-based pricing for personal care products. We find that gender segmentation is ubiquitous, as more than 80% of products sold are gendered. Further, we show that segmentation involves product differentiation; there is little overlap in the formulations of men's and women's products within the same category. Using a national dataset of grocery, convenience, drugstore, and mass merchandiser sales, we demonstrate that this differentiation sustains large price differences for men's and women's products made by the same manufacturer. In an apples-to-apples comparison of women's and men's products with similar ingredients, however, we do not find evidence of a systematic price premium for women's goods: price differences are small, and the women's variant is less expensive in three out of five categories. Our findings are consistent with the ease of arbitrage in posted price markets where consumer packaged goods are sold. These results call into question the need for and efficacy of recently proposed and enacted pink tax legislation, which mandates price parity for substantially similar gendered products.

*We would like to thank Sara Drango, Malika Korganbekova, Vitalii Tubdenov, and Jason Zhao for excellent research assistance. Thanks also to Eric Anderson, Jean-Pierre Dubé, Günter Hitsch, Ayelet Israeli, Anja Lambrecht, Amalia Miller, Aviv Nevo, and Yesim Orhun, as well as seminar participants at Georgia Tech, Northwestern - Kellogg, Temple, Vrije Universitat, Yale, the 14th FTC Microeconomics Conference, the Virtual Quant Marketing Seminar, and the QME Rossi Seminar hosted by the University of Rochester for helpful feedback.

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"); (iii) marketing databases provided through the respective NielsenIQ and the Nielsen Datasets; (iv) data from Syndigo LLC; and (v) data from Label Insight, Inc. at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the Nielsen, Syndigo, and Label Insight data are those of the researchers and do not reflect the views of NielsenIQ, Nielsen, Syndigo or Label Insight. Neither NielsenIQ, Nielsen, Syndigo, nor Label Insight is responsible for, had any role in, or was involved in analyzing and preparing the results reported herein.

Contact: sarah.moshary@chicagobooth.edu; anna.tuchman@kellogg.northwestern.edu; nbhatia@cornerstone.com.

1 Introduction

Price segmentation is a commonplace strategy employed by firms to increase profits in markets with heterogeneous demand. Segmentation based on gender for consumer packaged goods (CPG) has recently come under fire for creating an alleged *pink tax*, whereby goods marketed toward women are more expensive than their counterparts marketed toward men. A pink tax is concerning because it would exacerbate well-documented gender inequality in the labor market. Investigative journalists and government agencies report that gender price differences in CPG occur most frequently in personal care categories, such as deodorant and razors, and peg price differences in this category at 13% (e.g., Bessendorf 2015; Consumer Reports 2010; Duffin 2018). 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, which bans pricing on the basis of gender, and CA legislators passed a similar bill, AB 1287, in 2022. Since 2015, Representative Jackie Speier has introduced the Pink Tax Repeal Act in Congress four times, with the goal of implementing a similar ban nationwide. Unfortunately, there is a dearth of evidence on the pink tax to guide legislative action.

The aim of this paper is to provide systematic evidence on gender-based targeting and pricing for personal care products. A first finding is that gendering is ubiquitous; more than 80% of product volume is gender-targeted. We hypothesize that gender-based pricing operates differently in the market for personal care products compared to other markets studied in the academic literature on gender discrimination, including the labor market (see Blau and Kahn 2017 for a recent review), automotive vehicles, repairs, and loans (e.g., Ayres 1991; Ayres and Siegelman 1995; Goldberg 1996; Scott Morton, Zettelmeyer, and Silva-Risso 2003; Busse, Israeli, and Zettelmeyer 2017; Ozturk, He, and Chintagunta 2022), and real estate (Goldsmith-Pinkham and Shue Forthcoming). These settings entail price negotiation, so that 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, CPG like 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 barring her from buying a men's product. Economic theory predicts that to prevent arbitrage in these markets, firms must differentiate products targeted at different genders in such a way that consumers self-select into the product designed for their group (Mussa and Rosen 1978). 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. Price discrimination of this sort can be profitable for firms if men and women have different demand for soap; in particular, this strategy successfully

segments consumers by gender if women (but not men) are willing to pay more for the pink versus blue version of the product. While color is one attribute that a firm could use to segment consumers, differential demand for other product attributes could lead to product differentiation on other dimensions as well. For example, a soap manufacturer might choose to add shea butter, which has moisturizing properties, to its pink bar of soap if women have a higher willingness to pay for that ingredient than men.[1] Gender-based price discrimination in CPG would therefore manifest as differences in price across differentiated products that target different genders, rather than differences in the prices charged to men and women for the exact same product.

Accordingly, we find that differentiation is the rule rather than the exception among gendered personal care products, and this differentiation extends beyond product color and packaging. There is very limited overlap in the leading ingredients of men's and women's products in all six categories that we study: bar soap, body wash, deodorant, hair coloring, shampoo, and shaving cream. We view this product differentiation as integral to gender-based pricing in CPG, which we argue is a form of second degree price segmentation. Our viewpoint is at odds with the way that proposed and adopted legislation conceives of the pink tax; legislation bans price differences only in instances where a manufacturer sells "substantially similar" products targeted at men and women. The Pink Tax Repeal Act, the principal piece of proposed federal pink tax legislation, specifically references the materials used in the product and the product's intended use as criteria for evaluating similarity. Focusing on these apples-to-apples comparisons is akin to searching for instances of third degree price discrimination, cases where the prices of the men's and women's product variants differ but there is no substantial difference in horizontal or vertical quality between the variants.[2]

We hypothesize that the market should not sustain meaningfully different prices for men's and women's products that consumers truly perceive as substantially similar. We argue that if such price differences did exist, consumers would engage in arbitrage, rendering the strategy unprofitable. Of course, if a consumer faces a friction in purchasing a product targeted at another gender, then a firm may be able to sell substantially similar products to men and women at dissimilar prices. For example, if there is a social norm that pressures consumers to purchase and/or use products that align with their gender presentation, a firm could segment consumers simply by adding gender labels or color cues to their products. Alternatively, if consumers are uninformed or misinformed about product ingredients, they could mistake the differences between men's and women's products. We believe that this

[1] This example is motivated by our finding that astearic acid, a major component of shea and cocoa butter, is more prevalent in bar soaps targeted to women relative to those targeted to men (see Table A3 in Web Appendix C).

[2] The Pink Tax Repeal Act specifically rules out color as a meaningful differentiator (with an exception in cases where color generates cost disparities). This exception suggests that policymakers believe that a product's color does not confer value to consumers, yet can costlessly and effectively segment the market.

kind of widespread misperception is unlikely because information about ingredients is readily available on product packaging. Further, we find that married women buy the lion's share of personal care products for their households, including products targeted at both men and women, which suggests that they have some knowledge about the assortment of men's products. For example, 45% of married women buy men's deodorant on shopping trips where their husband is not present.[3] We acknowledge that in markets where this type of separation between the purchaser and the user of the product is not possible (e.g., haircuts), the direct interaction between seller and buyer could give greater scope for gender-based price discrimination. We hope future work will explore the pricing of these types of services, which are also subject to the Pink Tax Repeal Act.

We provide empirical evidence on the form and magnitude of gender-based price discrimination in personal care using Nielsen RMS data, which encompasses data on a wide array of products sold at thousands of retail outlets across the United States. We begin by estimating the average price difference between men's and women's products produced by the same manufacturer. We term this measure the *pink gap* because it can reflect both differences in markups and costs between the products that a manufacturer targets at men and women. Consistent with price discrimination, we find economically and statistically significant price differences within manufacturer. These price differences also tend to cut against women. Unit prices (e.g., price/oz or price/count) for women's products are higher than those for men's products in four of the nine categories we study. The remaining five categories do not have statistically significant unit price differences. Averaged across categories, the pink gap is 10.6%.

We next narrow the comparison to substantially similar products as per the Pink Tax Repeal Act, where we operationalize 'substantial similarity' as products made by the same manufacturer that contain the same leading ingredients. When we consider within-manufacturer comparisons of products with similar ingredients, price differences shrink and even reverse sign, so that in several categories men's products are more expensive. To obtain an overall estimate of the pink tax as defined by regulators, we pool the apples-to-apples estimates across categories and find that unit prices for women's products are 0.1% higher than for men's products. Together, the results support our hypothesis that firms engage in second rather than third degree price discrimination: we find evidence of meaningful price variation across differentiated products, but little price differentiation when comparing substantially similar products. One implication of our findings is 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. Further, these similar products are relatively few, limiting the applicability of the Pink Tax Repeal Act.

[3]This statement is based on an analysis of Numerator household panel data. The data record deodorant purchases made by 124,760 heterosexual married households in 2017 and the gender(s) of the shopper(s) on each trip.

Our work relates to several literatures in marketing and economics. First, our paper contributes to the extensive literature on price discrimination. Stigler (1987) defines price discrimination as "present when two or more similar goods are sold at prices that are in different ratios to marginal costs" (Varian 1989). Recent empirical studies of price discrimination in the marketing literature primarily focus on third-degree price discrimination or second-degree discrimination in the form of quantity discounts (Chintagunta, Dubé, and Singh 2003; Khan and Jain 2005; Dubé and Misra 2022). We argue that although regulators appear to conceptualize the pink tax as a form of third-degree price discrimination, gender-based targeting in the personal care category is nearer second-degree price discrimination; personal care products with differentiated features are targeted to different genders and sold at different prices (Mussa and Rosen 1978).

Second, we add to a large body of work exploring gender-based discrimination. The most notable and mature literature on gender discrimination studies the labor market (see Blau and Kahn (2017) for a review of this literature). A more recent but growing body of work explores gender discrimination in marketing contexts. For example, in the realm of digital advertising, Lambrecht and Tucker (2019) find that advertisers have greater demand for women's impressions, leading to higher prices for female eyeballs and divergence in the advertising content seen by women and men. Work by Busse, Israeli, and Zettelmeyer (2017) and Ozturk, He, and Chintagunta (2022) finds that women are charged higher prices for auto repairs and auto loans than men. They argue that price differences reflect statistical discrimination because gender correlates with consumer knowledge of market prices in these categories. We contribute to this literature by documenting the extent to which CPG are targeted by gender and whether such targeting leads to differential pricing. In our setting, the profitability of gender-based price discrimination depends on the extent to which gender identity correlates with preferences. In this way, our work ties to the consumer behavior literature that studies gender differences in preferences and responsiveness to advertising (Meyers-Levy and Loken 2015; Govind, Garg, and Mittal 2020; Gao, Mittal, and Zhang 2020; Fisher and Dube 2005; Dahl, Sengupta, and Vohs 2009).

Our study of the pink tax also relates to concepts of equity and fairness in the social choice literature and recent public policy debates and initiatives around diversity, equity, and inclusion. The social choice literature considers an allocation to be equitable if no agent prefers another agent's bundle to their own and fair if allocations are both equitable and pareto efficient (Varian 1975; Feldman and Kirman 1974). In the posted price markets that we study, consumers choose whether to buy a product that is targeted to men or women. Thus, absent an additional friction in the marketplace, consumer choices should be equitable per Varian (1975) and Feldman and Kirman (1974)'s definition of fairness. Nonetheless, fairness is a central concern in press coverage of the pink tax, suggesting that public opinion

employs a different criterion.[4] In terms of regulation, the national Pink Tax Repeal Act proposes that charging different prices based on gender for substantially similar products should be considered a violation of the Federal Trade Commission Act's Section 5, which covers unfair or deceptive practices in interstate commerce. One of the three factors the Commission explores to assess fairness is unjustified and substantial injury to consumers. The Commission clarifies that some injury may be acceptable if it is outweighed by any offsetting consumer or competitive benefits also produced by the practice (Pertschuk et al. 1980). In the personal care context, a benefit of second-degree price discrimination could include large assortments that are valued by heterogeneous consumers. Finally, the FTC states that the injury must be one which consumers cannot reasonably avoid. Our overall assessment is that (1) charging different prices for differentiated goods would not be considerd unfair by the Commission if the differentiation benefits some consumers; and (2) charging different prices for similar goods targeted to men and women would only be considered unfair if market frictions prevented arbitrage. As an example, the Commision may consider differential pricing unfair if it is sustained through marketing practices that exaggerate the differences between similar men's and women's goods. While we do not explore practices such as advertising that could facilitate price discrimination, we see this as a ripe and exciting opportunity for future work. Our focus in this paper is on the fundamental question of whether and to what extent firms charge different prices for similar goods targeted to men and women.

The paper proceeds as follows: Section 2 describes current pink tax legislation and the related literature, and Section 3 details the data. Section 4 provides evidence on product differentiation. Section 5 describes our preferred estimates of the pink tax. Section 6 discusses policy implications.

2 Current Legislation and Related Literature

The Pink Tax Repeal Act is the principal federal legislation aimed at combating price discrimination against women in CPG. Introduced by Congresswoman Jackie Speier, the act

> "prohibit[s] the sale of consumer products and services that are substantially similar if such products or services are priced differently based on the gender of the individuals for whose use the products are intended or marketed or for whom services are performed or offered."[5]

In practice, the Pink Tax Repeal Act considers products from the same manufacturer sub-

[4] E.g., Light (2022) and Mann (2018).
[5] The Pink Tax Repeal Act. H.R.3853. Accessed on November 12, 2021.

stantially similar "if there are no substantial differences in the materials used in the product, the intended use of the product, and the functional design and features of the product." 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.[6]

At the state level, both California and New York have passed 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, and in 2022, California legislators extended the existing restrictions on gender-based pricing to cover the pricing of goods as well. Similarly, in New York, a provision in the FY 2021 Budget Bill banned the pink tax. Like the Pink Tax Repeal Act, the CA and NY bills define the pink tax as gender-based pricing for substantially similar products, and they describe exemptions when differences in prices reflect differences in costs.[7]

To motivate these bills, Congresswoman Speier and other advocates for intervention frequently cite a report by the NYC Department of Consumer Affairs (Bessendorf 2015) that found substantial differences in 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.[8] We replicate and extend estimates from this study in Web Appendix G. Other studies of the pink tax have found more mixed results. Duesterhaus et al. (2011), which analyzes hand-collected prices from brick-and-mortar stores, found no significant difference in unit prices in three of the four personal care categories they study, but their analysis is likely under-powered. Contemporaneous work by Gonzalez Guittar et al. (2021) finds mixed results in their study of scraped price data from four online retailers, and a 2018 study by the Government Accountability Office finds that average prices are higher for women's products in five of ten personal care categories and men's products are more expensive in two.[9]

However, there is a conceptual disconnect between pink tax legislation and the way that these existing studies think about product attributes and price differences. The literature largely compares prices unconditional on product attributes. In contrast, pink tax legislation focuses on price differences between substantially similar products, suggesting that regulators conceptualize the pink tax as a form of third degree price discrimination.[10][11]

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

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

[8] 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.

[9] https://www.gao.gov/products/gao-18-500.

[10] Cohen, Elmachtoub, and Lei (2022) also describe pink tax legislation as banning third degree price discrimination based on gender.

[11] Such legislation is poised to be most meaningful if gendering restricts consumer choice to products targeted at a consumer's own gender, but gendering confers no additional value to consumers. For example, if women are restricted to buy the high-priced pink bar of soap even though the less expensive blue bar would provide them the same utility.

Drawing on economic theory, we argue that, absent some friction, the market for personal care products should not sustain meaningful price differences in instances where there are substantially similar products targeted at men and women. Instead, we contend that gendering in personal care is a second degree price discrimination phenomenon, one involving substantial differences in the attributes of products targeted at men and women. Further, we expect that product differentiation creates differences in both marginal cost and markups across gendered products. Accordingly, a central aim of this paper is to shed light on how gender-based pricing functions in CPG.

To distinguish between second and third degree price discrimination, we measure price differences for gendered products both conditional and unconditional on product ingredients. To our knowledge, the only other study that compares product ingredients and estimates prices accounting for product attributes is Wehner, Nead, and Lipoff (2017), a small-scale study of hair loss products that finds a 30% pink tax in a comparison of 16 pairs of gendered minoxidil medications with identical ingredients.

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,[12] razor blades, disposable and non-disposable razors, shampoo, and shaving cream. The data record the price and quantity sold for products (UPCs) sold in 39,697 stores affiliated with 93 chains across the US. The raw data are recorded at the store-UPC-week level, so that prices reflect the weekly average price paid by consumers in a particular store and week. The data also include product characteristics such as brand name and product size.

The data do not indicate the price of a product in weeks when it earns no sales at that store. We must therefore impute prices for our analysis. 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.

After imputing prices in weeks with no purchase, we aggregate the Nielsen data to the store-UPC-year level by summing quantities and taking a simple average of price over weeks

[12] 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.

in the year.[13] The temporal aggregation helps with computational feasibility — even after aggregating the data to the yearly level, we have upwards of 30 million observations in the deodorant category alone — while still allowing us to compare the average prices for products targeted to men and women. We keep the data at the store level rather than aggregating to the chain level because assortments may vary across stores in the same chain, and pink tax legislation centers on the comparison of prices within a store.

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.[14]

2. `Walgreens.com`**:** We scraped gender categorizations from the website of the large American drugstore chain in Summer 2020. Web Appendix B.1 displays screenshots of gender categorization and filters on the website.

3. **Differential purchasing by all-male and all-female households in the Nielsen Consumer Panel dataset:** 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.[15] 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%). 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.

4. **Label Insight:** We collect data on gender targeting from Label Insight, a market research firm that records marketing claims for CPG brands.[16] The database also includes product pictures.

[13] We take the simple average instead of a quantity-weighted average because our focus is on price charged, as opposed to price paid.

[14] The following words were used to classify men's UPCs: his, men, clubman, hombre, man, homme, men's, monsieur, and Mr. The following words and abbreviations were used to classify women's UPCs: her, lady, girl, ldy, women, femme, ladies, lady's, and wmn.

[15] Women-only households represent 71% of single-gendered households.

[16] The gender field is populated for a subset of the personal care products.

5. **Hand-coding Label Insight product images:** We hired undergraduate research assistants to categorize product images from Label Insight as male, female, unisex, or unknown. Web Appendix B.2 provides more 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. In a final step, we fill in gender targeting 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: Gender Targeting across Personal Care Categories

Nielsen Product Module	(1)	(2)	(3)	(4)	(5)	(6)
	% Qty Gendered		Total	% UPCs Gendered		Count of
	All	for Women	Qty (MM)	All	for Women	UPCs
Soap - Bar	72.0%	57.4%	523	20.6%	56.6%	3,710
Soap - Liquid	46.3%	94.2%	508	13.8%	91.1%	3,100
Soap - Specialty	78.5%	63.1%	820	28.0%	61.5%	6,889
Deodorants - Personal	99.0%	48.8%	1,061	71.8%	41.3%	2,965
Hair Coloring	100.0%	88.9%	312	100.0%	95.8%	2,534
Hand & Body Lotions	73.3%	95.5%	472	20.5%	91.5%	7,593
Razor Blades	87.9%	33.3%	96	60.5%	36.9%	519
Razors Disposable	69.4%	51.1%	289	42.2%	49.6%	978
Razors Non-Disposable	90.0%	46.1%	68	57.4%	39.0%	451
Creme Rinses & Conditioners	83.3%	99.7%	551	32.2%	96.4%	5,916
Shampoo						
–Aerosol/ Liquid/ Lotion/ Powder	84.8%	80.1%	895	41.0%	79.7%	7,746
–Bars/ Concentrates/ And Creams	68.6%	89.6%	9	41.2%	90.7%	131
–Combinations	55.6%	99.9%	20	27.7%	88.9%	519
Shaving Cream	100.0%	25.2%	271	100.0%	21.9%	942
Overall	81.7%	68.0%	5,895	37.6%	72.0%	43,993

Notes: This table describes the share of products available at Nielsen RMS stores between 2015-2018 that we record as gendered.

Table 1 shows the pervasiveness of gender targeting across product modules.[17] Consistent with the focus on personal care in the media surrounding the pink tax, we find that personal care categories are highly gendered. Column (1) shows that our methods assign a gender to the lion's share of category volume sales (82%). We consider this an underestimate of gender-targeting in personal care because our methods have a hard time identifying targeting for niche products. This is reflected in the lower share (38%) of categorized personal care UPCs in column (4). We also find considerable variation across categories in the market share of gendered products and in the relative prevalence of products targeted at men vs. women; for

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

example, column (2) shows that the overwhelming share of gendered hair coloring products are targeted at women, but most gendered shaving creams are targeted at men.[18]

These patterns naturally raise a question surrounding the availability and pricing of unisex products. As shown in Table 1, for the personal care categories that we study, gender-targeting is the rule rather than the exception. This finding is consistent with the focus on these products in the press surrounding the pink tax. Thus, we think of unisex products as niche products in this setting. Unfortunately, our data is not well-suited to studying unisex products because they are difficult to identify. The challenge is that when our methods fail to assign a gender label to a product in the Nielsen RMS dataset, we cannot tell whether the product is not gender-targeted (unisex), or if the product truly is gender-targeted but we are missing that information. This ambiguity arises naturally for niche products for which there are too few recorded household purchases in the Nielsen HMS data to infer a clear gender association (method 3). Niche products are also less likely to be carried by Walgreens.com (method 2) and be included in the Label Insight data (methods 4 and 5). However, among products with explicit gender information on Walgreens.com and in the personal care categories that we study, we find that only 6.6% are labeled "unisex," providing further evidence that gender-targeting is the norm.

Our data also sheds light on the way that firms signal gender-targeting to customers. We asked three of the research assistants who assessed gender-targeting from product images to indicate which feature(s) of the image indicated gender. Color was the most commonly cited attribute, with at least one research assistant indicating color as a means of identifying the gender target for 98% of product images from Label Insight.[19] The product name and description was the second most-cited attribute that indicated gender. These patterns highlight that the manufacturer plays a crucial role in gender targeting through the naming and packaging of its products. We therefore think of gender-based segmentation as a manufacturer-driven phenomenon. Of course, retailers can reinforce segmentation through search filters on their websites (as we find for `Walgreens.com`) or by assorting women's and men's products in different aisles of their brick-and-mortar stores. In our main empirical specification, we examine price differences conditional on both manufacturer and retailer.

Ingredients

We use Syndigo data on product ingredients to assess the similarity of different products. For each UPC, the data include the names and amounts of any active ingredients, as well as

[18] Thirty-one percent of disposable razors sold are not assigned a gender. This relatively high share may be driven by private label products (27% market share, as shown in Web Appendix B.4). Because Nielsen masks the UPCs of private label products, we cannot map them to our data sources for gender.

[19] Three research assistants were asked to indicate their reason for determining that a product is targeted at men or women for 962 images, all in the deodorant category.

a list of inactive ingredients. The data cover about 85% of the sales of personal care products in our sample of gendered products from the Nielsen data. Web Appendix C reports the shares by category. Table 2 reports the median number of inactive ingredients by category, which range from 10 (deodorants) to 55 (hair coloring). When we construct comparisons of substantially similar products, we define products as similar if they contain the same first few leading ingredients.[20] Because most products in our sample have lengthy ingredient lists, we consider this approach to be conservative, admitting comparisons between products that consumers might perceive as dissimilar. We discuss this further in Section 5.

Table 2: Median Number of Product Ingredients by Category

Module	(1) Overall	(2) Men	(3) Women
Bar Soap	16	15	17
Body Wash	18	15	19
Deodorants	10	10	11
Hair Coloring	55	10	56
Shampoo	22	19	23
Shaving Cream	18	19	18
Total	18	14	21

Notes: This table reports the median number of inactive product ingredients by category. Columns (2) & (3) report the median number for men's and women's products separately.

Documenting Price Patterns

We begin by describing the overall distribution of prices for personal care products targeted at men and women. Panel (1) of Table 3 reports the average shelf price of products in our sample separately by category and gender. Men's products are more expensive in six of nine categories. 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.

However, we also find that women's products tend to be smaller than men's products, so lower product shelf prices might obfuscate higher per-unit prices. To address this concern, panel (2) reports average unit shelf prices (i.e., price per ounce or count, depending on the category). On a per-unit basis, women's products are more expensive than men's products in five of nine categories. While men's products are more expensive in the other four categories, we note that the price differences are largest in the categories where the prices of women's products are higher.

[20] Table A3 in Web Appendix C summarizes the ingredients that most frequently appear in the top five, separately by product category and gender.

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

These descriptive statistics are instructive about the difference in the average prices of product assortments targeted at men and women, but they do not speak directly to gender-based price discrimination. First, these price differences do not control for product attributes (other than size). Understanding differences in the ingredients of men's and women's products is the focus of the next section. Second, the price differences in Table 3 incorporate differences in the pricing of men's vs women's products both within and across manufacturers. As shown in panel (3), while many manufacturers sell both men's and women's products, a substantial number specialize in one or the other. The manufacturers that specialize tend to have smaller market shares, as reported in panel (4). As per pink tax legislation, our analyses will focus on the larger manufacturers who cater to both men and women through separate gendered product lines.

Table 3: Summary Statistics on Price and Manufacturers

	(1)		(2)		(3)		(4)	
	Price per Product		Price per Unit		N Manufacturers		Share Manufacturers	
Module	Men	Women	Men	Women	Both	Only One	Both	Only One
Bar Soap	4.31	4.66	0.23	0.48	11	38	91.1%	8.9%
Body Wash	4.47	5.28	0.28	0.51	16	50	90.7%	9.3%
Deodorant	4.59	4.93	1.51	2.02	25	20	98.6%	1.4%
Hair Coloring	8.65	7.88	8.40	7.87	6	16	74.9%	25.1%
Razor Blades	21.50	17.82	3.59	4.23	3	5	93.2%	6.8%
Razors Disposable	7.67	6.96	2.33	2.02	5	4	96.0%	4.0%
Razors Non-Disposable	11.44	10.33	11.44	10.33	4	7	91.3%	8.7%
Shampoo	6.30	5.64	0.49	0.52	25	78	84.8%	15.2%
Shaving Cream	3.68	3.13	0.62	0.48	14	47	94.7%	5.3%

Notes: This table reports summary statistics broken down by category and gender. Panel (1) reports the average price in dollars of men's and women's products. Panel (2) reports the average price per unit of measure (oz or count) in dollars of men's and women's products. Panel (3) reports the number of manufacturers that produce products for both genders vs. those that produce for only one of the genders. Panel (4) reports the market share in terms of unit sales of manufacturers that produce products for both genders vs. for only one of the genders.

4 Product Differentiation

This section examines the comparability of men's and women's personal care products. While policy-makers are focused on eliminating price differences among substantially similar products targeted at men and women, economic theory suggests that the market should not sustain meaningful price differences for men's and women's products that are truly similar. Facing a choice between a men's product and a women's product that are in essence the same,

canonical models of utility-maximization imply that a consumer should purchase the cheapest version of the product. If consumers behave in this fashion, then in order to successfully segment the market, a firm must second degree price discriminate via product differentiation that induces consumers to voluntarily sort into their own gender's target product.

We turn to the data to test whether firms that engage in gender segmentation do so via product differentiation. The Pink Tax Repeal Act defines products as substantially similar if "there are no substantial differences in the materials used in the product, the intended use of the product, and the functional design and features of the product." Following these guidelines, we assess similarity of products within the same category on the basis of product ingredients. We define a product formulation as the combination of manufacturer, active ingredient, and the first five inactive ingredients, where the order of ingredients matters.[22] Table A3 in Web Appendix C lists the most common leading ingredients for men's and women's products by category. In using a subset of product ingredients, our goal is to identify products that are similar without requiring products to be identical. This strategy is in keeping with the spirit of how the Pink Tax Repeal Act deals with product color; it does not consider color to be a meaningful differentiator and so bars price differences for products that differ in color but are otherwise identical. Because the law also explicitly restricts to within-store price comparisons, we restrict comparisons to products sold at the same retail outlet.

Table 4 describes the extent of overlap in manufacturer-formulations across genders within store.[23] As an example, column (1) of Table 4 shows that the average store carries 40 unique deodorant formulations targeted to women and 41 targeted to men. Column (2) shows that approximately 32% of formulations targeted to women have a comparable formulation targeted to men within the same store and vice versa. Formulation overlap varies across categories: hair coloring products have almost no overlap across genders; deodorants have the most overlap. Across personal care, less than half of gendered UPCs in the average store have a comparable product targeted at the other gender (see column (4)). Column (6) reports the market share captured by products with overlapping formulations. In some categories (e.g., body wash) products with a formulaic analog are more popular than those without, while in other categories (e.g., deodorants) the non-overlapping formulations are more popular.

The patterns in Table 4 offer two implications. First, they confirm that manufacturers differentiate products targeted at men and women through ingredients. On a given shopping

[22]The FDA requires active ingredients be reported first, then inactive ingredients in order of predominance. Any order is permitted for inactive ingredients comprising less than 1% of the product. [https://www.fda.gov/cosmetics/ cosmetics-labeling-regulations/cosmetics-labeling-guide#clgl]. We also explore the robustness of our results to an alternative definition of formulation in which the order of the top ingredients does not matter.

[23]We use data from 2018 to ensure similar formulations were available at the same point in time. We exclude convenience stores because they carry very limited assortments. Table A4 in Web Appendix C shows results using the first three inactive ingredients to define a product formulation.

Table 4: 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	10	15.9%	20	18.3%	1,040	14.7%
Bar Soap	women	12	13.5%	29	14.2%	1,500	17.0%
Body Wash	men	24	26.6%	44	28.3%	1,658	30.3%
Body Wash	women	55	12.6%	87	19.4%	2,757	25.5%
Deodorants	men	41	32.0%	116	39.7%	4,069	32.6%
Deodorants	women	40	32.7%	103	55.6%	3,714	51.0%
Hair Coloring	men	5	0.0%	15	0.0%	243	0.0%
Hair Coloring	women	16	0.0%	167	0.0%	1,753	0.0%
Shampoo	men	18	23.5%	35	28.1%	961	27.5%
Shampoo	women	79	5.1%	132	11.1%	3,356	13.9%
Shaving Cream	men	15	17.6%	31	23.4%	1,360	23.4%
Shaving Cream	women	5	48.5%	10	54.8%	346	61.9%

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 ingredient information in the Syndigo data. Convenience stores are excluded because they have very small assortments.

trip, a consumer examining a women's personal care product at random would not find an analog targeted at men in the same store. Rather, the men's products produced by the same manufacturer would differ not only in the color of the package or the product label, but also in the main ingredients that comprise the product's formulation. The dissimilarity in product formulation across gendered products constitutes an important piece of evidence on how price discrimination functions for personal care products. Differentiation is necessary when a firm engages in second degree targeting. In contrast, a firm that can third degree discriminate need not differentiate its offerings to different segments. A second implication of these patterns is that legislation like the Pink Tax Repeal Act would not apply to most personal care products because a substantially similar analog does not exist for most products at most stores.

Finally, we note that firms may differentiate products on dimensions beyond ingredients. The differences in per product and per unit prices in Table 3 suggest that package size is one such dimension. In Web Appendix D, we confirm that products targeted to men tend to be larger, even when we restrict to comparisons of men's and women's products made by the same manufacturer. These size differences raise a question of whether quantity discounts could be confounded with a pink tax, a possibility that we will return to in our analysis of prices below.

5 Measuring Price Disparities

This section presents evidence on the extent to which firms engage in gender-based price discrimination. We begin by estimating the difference in the average price of men's vs women's products that are produced by the same manufacturer and sold at the same retail outlet. We term this difference the *pink gap* because it does not control for differences in the ingredients of men's and women's products. Thus, the pink gap can reflect both second and third degree price discrimination.

Our main specification models the price of product j sold by store s in year t, p_{jst}, as a function of its intended gender target, manufacturer fixed effects (γ_m), year fixed effects (γ_t), and store fixed effects (γ_s).[24] Including manufacturer fixed effects is key to identifying price discrimination, where a single firm sells to consumers in different segments at different prices. The store and year 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

$$p_{jst} = \beta \cdot women_j + \gamma_m + \gamma_t + \gamma_s + \varepsilon_{jst}, \tag{1}$$

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. Because men's and women's products tend to differ in size (see Web Appendix D), we analyze price per unit of measure (price/oz or price/count) as our dependent variable. In this way, β measures differences in the prices that are charged for a comparable amount of product. We estimate equation (1) separately for each personal care category.[25] Standard errors are two-way clustered at the store and year level.

The results are reported in column (1) of Table 5. On a per unit basis, women's products are more expensive in four of nine categories. The remaining five categories do not have statistically significant differences in unit prices. Furthermore, the magnitudes of the percentage price gaps are fairly large. Taking a simple average across all nine categories, women's personal care products are 10.6% more expensive than men's personal care products. These results indicate that manufacturers do engage in some form of gender-based price discrimination.

Next, we aim to distinguish whether the pink gap we document in column (1) reflects second or third degree price discrimination. With this goal in mind, we augment our regression specification to isolate comparisons between otherwise similar products that are targeted

[24] Observations are weighted by the number of weeks that the product is on a store shelf. Web Appendix A describes our procedure for imputing whether a product is available in a store in a particular week.

[25] We analyze 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 and that have at least 10 million units sold.

Table 5: Price Gap by Category, 2015-2018

Module	(1) Unit Shelf Price	(2) Unit Shelf Price	(3) Unit Shelf Price	(4) Unit Shelf Price
Bar Soap	0.08***	0.09***	0.03***	0.04***
	(0.00)	(0.00)	(0.00)	(0.00)
	0.23	0.23	0.23	0.23
	35.4%	37.6%	14.9%	15.9%
	7,832,795	7,273,999	7,273,999	7,273,999
Body Wash	0.02*	0.02*	-0.02***	-0.01**
	(0.01)	(0.01)	(0.00)	(0.00)
	0.28	0.28	0.28	0.28
	8.7%	6.7%	-8.4%	-5.1%
	20,076,939	18,264,364	18,264,364	18,264,364
Deodorant	0.45***	0.44***	0.20**	0.06
	(0.02)	(0.01)	(0.05)	(0.06)
	1.50	1.51	1.51	1.51
	29.6%	29.0%	13.5%	4.3%
	31,001,944	29,727,058	29,727,058	29,727,058
Hair Coloring	0.55*			
	(0.20)			
	8.49			
	6.5%			
	28,450,230			
Razor Blades	0.76***			
	(0.04)			
	3.68			
	20.7%			
	3,997,551			
Razors Disposable	-0.05			
	(0.09)			
	2.28			
	-2.3%			
	6,014,671			
Razors Non-Disposable	-0.27			
	(0.18)			
	11.69			
	-2.3%			
	3,158,478			
Shampoo	0.01**	0.00	0.00	-0.03***
	(0.00)	(0.00)	(0.01)	(0.00)
	0.50	0.51	0.51	0.51
	2.9%	0.2%	-0.5%	-6.5%
	30,835,605	23,296,620	23,296,620	23,296,620
Shaving Cream	-0.02	-0.02*	-0.04***	-0.04***
	(0.01)	(0.01)	(0.00)	(0.00)
	0.59	0.57	0.57	0.57
	-3.6%	-3.6%	-7.2%	-7.9%
	6,773,794	5,746,248	5,746,248	5,746,248
Data	All	Syndigo	Syndigo	Syndigo
Manufacturer FE	Y	Y	N	N
Formulation FE	N	N	Top 3	Top 5

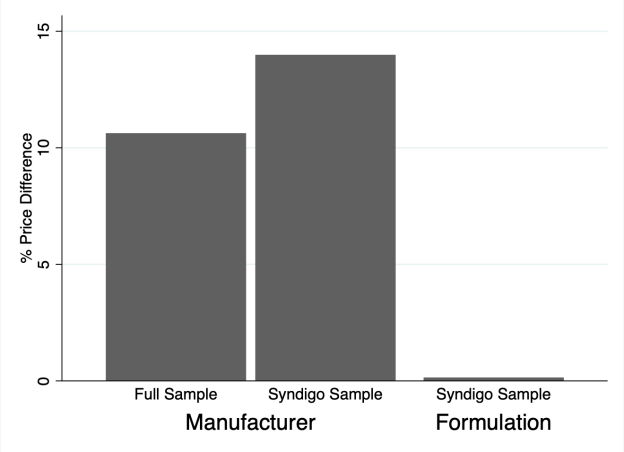
Notes: The sample in column (1) comprises the full set of products. Columns (2)-(4) exclude razors because their ingredients are not reported and hair coloring because there is insufficient overlap in ingredients across men's and women's products. Columns (3) and (4) include formulation fixed effects. For each category, the first row reports the average price difference and the second row reports the standard error (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 difference, calculated as the ratio of row one to row three. The fifth row reports the number of observations. Regressions are estimated separately by product module and include store and year fixed effects.*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

to men vs women by controlling for product ingredients. These estimates that control for store, manufacturer, and product formulations are our preferred measure of the pink tax as conceptualized by policy-makers. This analysis subsets to the sample of Nielsen products that merge to Syndigo data on product ingredients. While we do not observe ingredient information for all products in the Nielsen data, the products for which we observe ingredients account for the majority (85%) of sales in the Nielsen data (see Web Appendix Table A2 for the break-down by category). Further, we find that the merged sample is similar to the full set of Nielsen products in terms of the unconditional pink gap, as shown in column (2) in Table 5, which replicates the regression in column (1) for products with ingredient information. Column (3) displays price differences when we add fixed effects for product formulation, which we define as a unique combination of manufacturer, active ingredient, and the first three inactive ingredients. Column (4) expands the analysis to include the top five inactive ingredients in the definition of product formulation. The gender price gap shrinks towards zero in three of the five product categories when we control for manufacturer and leading ingredients. Women also face lower prices in three of the five product categories: body wash, shampoo, and shaving cream. Women face higher prices for bar soap, and prices are roughly comparable for deodorants.

Because the courts have not clarified what constitutes a 'substantial difference in materials,' our consideration of the first three to five inactive ingredients is admittedly somewhat arbitrary. The median product in our sample comprises 19 ingredients, so it seems plausible that this criterion omits key ingredients for some products. Our specification may therefore admit comparisons between products that are in fact substantially different; in this case, we would tend to falsely attribute differences in costs to differences in price. On the other hand, if some combinations of ingredients are in fact functionally similar, our set of comparisons could be overly restrictive, obscuring instances of price discrimination. To address this concern, we explore the robustness of our results to an alternative definition of formulation in which the order of the top ingredients does not matter (see Web Appendix C). The results are very similar, largely because there are few men's and women's products that have the same top three or top five ingredients just in a different order.

We now return to the discussion of product size and consider the role that quantity discounts play in driving a price differential. In Web Appendix D, we show that controlling for size in our price per oz regressions tends to shift our pink gap estimates down. In some cases, the sign of our pink gap estimate even flips from positive to negative. This analysis reveals that quantity discounts are economically meaningful for personal care products, but also that these discounts are *differentially offered* for men's products. We see the estimates unconditional on product size as best aligned with how regulators and consumers conceptualize a pink gap/tax because these estimates capture the price differences that consumers face

Figure 1: Average Price Differences across Categories



Notes: This figure displays the average (unweighted) percentage price difference between women's and men's products across categories. The label "Manufacturer" indicates that the estimate is based on a specification with manufacturer fixed effects. The label "Formulation" indicates that the underlying specification included fixed effects for each combination of manufacturer, active ingredient, and top five product ingredients.

for products targeted to men vs women. Seen in this light, the regressions without quantity controls measure the pink tax, while the regressions controlling for product size help explain the "mechanism" underlying the price differences.

To compute the overall price difference for personal care products, we average the estimated percentage price differences from Table 5 across categories. Figure 1 presents these averages. Within manufacturer, unit prices for women's products are 10.6% more expensive than unit prices for men's products.[26] For categories with ingredients data, the average pink gap is slightly larger, at 14.0%. However, when we condition on formulation, unit prices for women's products are on average 0.1% more expensive per unit.[27] Motivated by the recent adoption of pink tax regulations at the state level, Web Appendix F presents estimates of price differences separately by state. For all states, we find a positive unconditional gender price gap and a small, negative pink tax. Notably, our pink tax estimates are similar in states with and without enacted or pending pink tax legislation. Of course, these simple averages of percentage price differences do not account for differences in price levels and the relative frequency of purchase across product categories.

To address these considerations, we conduct a back-of-the-envelope calculation of how

[26] As a robustness check, we estimated the pink gap using data on prices from Walgreens.com. A feature of this data is that it does not require price imputation, although the sample size is small ($N = 595$). The estimated pink gap is positive and the corresponding 95% CI ranges from 10.3% to 29.7%, covering the pink gap estimate from the RMS data.

[27] We also consider an alternative weighting whereby we re-estimate equation (1) stacking the data across categories and using log price per unit of measure as the dependent variable. This analysis puts more weight on categories with more observations (i.e. those that have more products in more stores). Under this alternative weighting, we find that, within manufacturer, unit prices for women's products are 9.5% more expensive than men's products. In a comparison of products with comparable formulations, women's products are 2.5% less expensive.

much an average household could save by substituting to cheaper products targeted to a different gender.[28] We find 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 9%–if a household were willing instead to substitute to products with different gender targeting and different formulations. However, a revealed preference argument suggests that such switching would lower consumer welfare.

Taken together, the estimates in Table 5 do not support the hypothesis that women systematically face higher prices than men for comparable personal care products. Rather, they tell a more complex story: there are economically and statistically significant price differences across gendered products, and on average women's products are more expensive, but the difference in prices is driven by a comparison of differentiated products. Price differences shrink and even reverse sign 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

This paper shows that gender-based segmentation is pervasive in the market for personal care products, but we argue that it is not analogous to the gender discrimination documented in the labor market. In markets where prices are negotiated, firms can tailor the wage of a given job (or price of a given product) to the gender of the employee (or buyer). However, consumer packaged goods are sold in posted price markets, which prevents firms from quoting buyer-specific prices for a particular product. Rather, economic theory suggests that successful price discrimination by gender requires that firms differentiate products, offering variants that appeal to consumers of different genders. Using a national dataset of prices from grocery, convenience, drug, and mass merchandise retail outlets, coupled with detailed data on product gender targeting and ingredients, we show that differentiation is the rule rather than the exception among gendered personal care products; the same manufacturer often chooses different leading ingredients for its men's and women's products. An analysis of the purchases of single-person households in the Nielsen Consumer Panel reveals that this strategy is successful; we find that women primarily purchase items targeted at women, and vice versa for men. As an example, 78% of women and 81% of men exclusively purchase deodorants targeted at their own gender.[29]

[28] Web Appendix E.2 describes the analysis.

[29] The same patterns hold if we rely only on explicit gender-targeting that excludes the panelist-based categorization. We present results on the effectiveness of gender targeting in other modules in Web Appendix Table A8.

Turning to prices, we find meaningful differences in the retail prices of men's and women's products that are manufactured by the same firm. Women's products are not always more expensive: in five of the nine categories we study, the difference in price between men's and women's products is neither statistically nor economically significant. Our estimates peg the pink gap, which we define as the within-manufacturer difference in price between men's and women's products, at about 11%. A positive pink gap does not necessarily imply that women enjoy lower average consumer surplus than men in the market for personal care products—that will depend on how women and men value product quality. Indeed, in the classic second-degree price discrimination model, the consumer type with a higher taste for quality buys a more expensive, higher-quality product and enjoys higher consumer surplus (Mussa and Rosen, 1978; Bolton and Dewatripont, 2005).

Our view of gender-based pricing for personal care products contrasts with the current policy approach to regulating gender-based pricing in CPG. Proposed and enacted legislation seeks to ban the "pink tax," defined as any difference in the prices of substantially similar women's and men's products. We question how such price differences could persist in equilibrium because consumers should purchase the lowest priced alternative among the set of products that they truly consider similar, regardless of the product's intended gender target. The data provide support for our conjecture. In an apples-to-apples comparison of products with the same top ingredients that are produced by the same manufacturer, price differences between men's and women's products are economically small.

These findings on the pink tax accord with the relative ease of arbitrage across gendered products in posted price markets. However, our results contrast markedly with estimates from Bessendorf (2015), a New York City Department of Consumer Affairs (NYC DCA) report frequently referenced in pink tax legislation. Bessendorf (2015) report economically meaningful differences between the prices of substantially similar men's and women's products based on a hand-collected sample of prices for 122 personal care products at three NYC drugstores. Although the sample was constructed by subjectively pairing men's and women's products, we find that most pairs in the sample differ in their top ingredients and thus do not speak to the substantially similar comparisons outlined in the Pink Tax Repeal Act. This finding dovetails with our broader finding that products targeted at men and women typically differ on other dimensions besides color. We also note that the products considered in Bessendorf (2015) account for less than 6% of category sales and were not selected at random. Web Appendix G provides a deeper replication and evaluation of the analysis in the NYC DCA report. The divergence in our conclusions highlights the importance of bringing to bear comprehensive and representative data and systematic methods (e.g., for gender-coding) in empirical research.

Our findings provide several implications for current and proposed legislation. First, our

finding that women's personal care products are not systematically more expensive than substantially similar men's products calls into question the role of government intervention to reduce the pink tax. We acknowledge that our analysis has limitations: our findings do not speak to gender-based price discrimination in service industries (e.g., haircuts) and even for personal care, our findings pertain to average price differences, which may mask instances of a particular retail outlet or manufacturer pricing in a way precluded by the Pink Tax Repeal Act.[30] 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 average gender price gap, but the Pink Tax Repeal Act would require that both retailers change their pricing policy or product assortments. However, our analysis reveals that most women's products do not have a men's analog sold in the same 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 analog, the Pink Tax Repeal Act might induce the retailer to drop the men's variant, de facto increasing price dispersion by setting an infinite price. In this case, the Pink Tax Repeal Act could reduce consumer surplus by decreasing the total quanity sold.

Finally, while our findings do not support the existence of a pink tax as conceived by regulators, a more expansive definition of the pink tax could include differences in markups across men's and women's products. Put another way, the pink tax could be redefined to encompass second degree price discrimination. 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, firms 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. We hope to see more work exploring how and in what way demand varies by gender, which can help shed light on why firms segment on gender rather than other consumer characteristics, such as price sensitivity.

[30] We explored heterogeneity in the pink gap across manufacturers and did not detect any consistent outliers.

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.

Bolton, Patrick and Mathias Dewatripont (2005). *Contract Theory*. The MIT Press.

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.

Chintagunta, Pradeep, Jean-Pierre Dubé, and Vishal Singh (2003). "Balancing Profitability and Customer Welfare in a Supermarket Chain". In: *Quantitative Marketing and Economics* 1, pp. 111–147.

Cohen, Maxime C., Adam N. Elmachtoub, and Xiao Lei (2022). "Price Discrimination with Fairness Constraints". In: *Management Science*.

Consumer Reports (2010). *Men Win the Battle of the Sexes*. URL: https://www.consumerreports.org/cro/2012/05/men-win-the-battle-of-the-sexes/index.htm (visited on 2010).

Dahl, Darren W., Jaideep Sengupta, and Kathleen D. Vohs (2009). "Sex in Advertising: Gender Differences and the Role of Relationship Commitment". In: *Journal of Consumer Research* 36.2, pp. 215–231.

Dubé, Jean-Pierre and Sanjog Misra (2022). "Personalized Pricing and Consumer Welfare". In: *Journal of Political Economy*.

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 (2018). *The Problem with the Pink Tax*. (Visited on 12/06/2019).

Feldman, Allan and Alan Kirman (1974). "Fairness and Envy". In: *The American Economic Review* 64.6, pp. 995–1005.

Fisher, Robert J. and Laurette Dube (2005). "Gender Differences in Responses to Emotional Advertising: A Social Desirability Perspective". In: *Journal of Consumer Research* 31.4, pp. 850–858.

Gao, Huachao, Vikas Mittal, and Yinlong Zhang (2020). "The Differential Effect of Local-Global Identity Among Males and Females: The Case of Price Sensitivity". In: *Journal of Marketing Research* 57.1, pp. 173–191.

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 (Forthcoming). "The Gender Gap in Housing Returns". In: *Journal of Finance*.

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

Govind, Rahul, Nitika Garg, and Vikas Mittal (2020). "Weather, Affect, and Preference for Hedonic Products: The Moderating Role of Gender". In: *Journal of Marketing Research* 57.4, pp. 717–738.

Hitsch, Günter, Ali Hortaçsu, and Xiliang Lin (2021). "Prices and Promotions in U.S. Retail Markets". In: *Quantitative Marketing and Economics* 19, pp. 289–368.

Khan, Romana J. and Dipak C. Jain (2005). "An Empirical Analysis of Price Discrimination Mechanisms and Retailer Profitability". In: *Journal of Marketing Research* 42.4, pp. 516–524.

Lambrecht, Anja and Catherine Tucker (2019). "Algorithmic Bias? An Empirical Study of Apparent Gender-Based Discrimination in the Display of STEM Career Ads". In: *Management Science* 65.7, pp. 2947–3448.

There's a Pink Tax on Women (2022).

Feminists in Line for Free Lipstick (2018).

Meyers-Levy, Joan and Barbara Loken (2015). "Revisiting Gender Differences: What We Know and What Lies Ahead". In: *Journal of Consumer Psychology* 25.1, pp. 129–149.

Mussa, Michael and Sherwin Rosen (1978). "Monopoly and Product Quality". In: *Journal of Economic Theory* 18, pp. 301–317.

Ozturk, O. Cem, Cheng He, and Pradeep Chintagunta (2022). "Do Women Pay More Than Men for Indirect Auto Financing?" In: *Working Paper*.

Pertschuk, Michael et al. (1980). *FTC Policy Statement on Unfairness*. Tech. rep. Federal Trade Commission.

Scott Morton, Fiona, Florian Zettelmeyer, and Jorge Silva-Risso (2003). "Consumer Information and Price Discrimination: Does the Internet Affect the Pricing of New Cars to Women and Minorities?" In: *Quantitative Marketing and Economics* 1, pp. 65–92.

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

Stigler, George (1987). *Theory of Price*. Macmillan.

Varian, Hal R. (1975). "Two Problems in the Theory of Fairness". In: *Journal of Public Economics* 5.3-4, pp. 249–260.

— (1989). "Handbook of Industrial Organiziation". In: ed. by R. Schmalensee and R.D. Willig. Elsevier Science Publishers B.V. Chap. 10, pp. 597–654.

Wehner, Mackenzie R., Kevin T. Nead, and Jules B. Lipoff (2017). "Association Between Gender and Drug Cost for Over-the-Counter Minoxidil". In: *JAMA Dermatology* 153.8, pp. 825–826.

Web Appendix

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 observation that retailers rarely change their regular shelf price for a product, and instead create short-term variation in prices by running temporary price promotions as discounts off of 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 four 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.

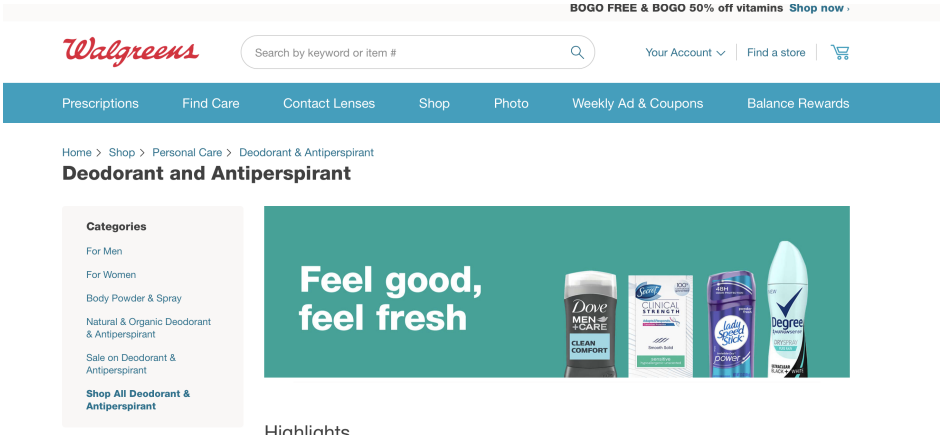
B Gender-Targeting Data Sources

B.1 Walgreens

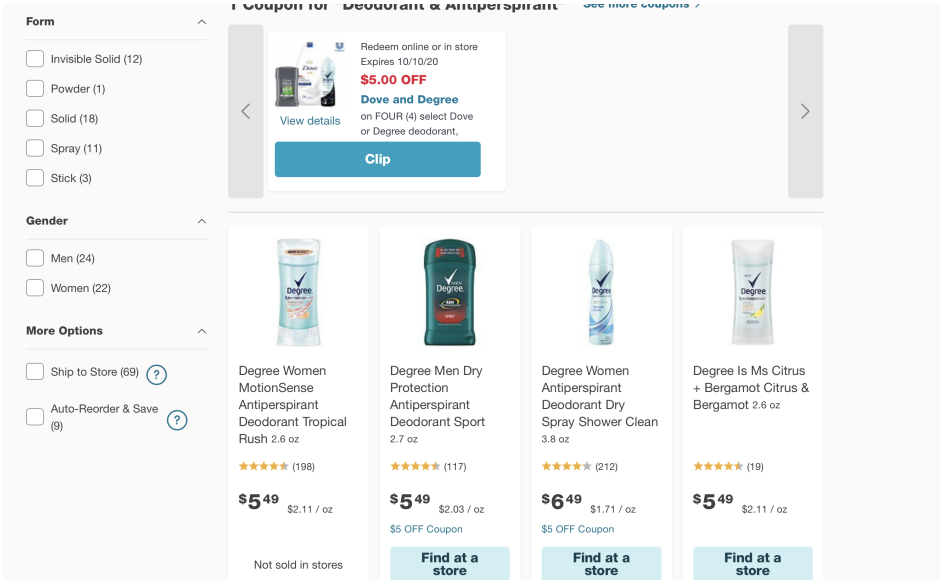
We extract gender information from the Walgreens website. The website explicitly categorizes certain product categories by gender. Figure A1 (a) presents one such example for the Deodorant & Antiperspirant category. We also collect gender information from search result page gender filters, as in Figure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

Figure A1: Walgreens.com Gender Categorizations

(a) Primary Gender Classification



(b) Gender Filter on Search Results Page



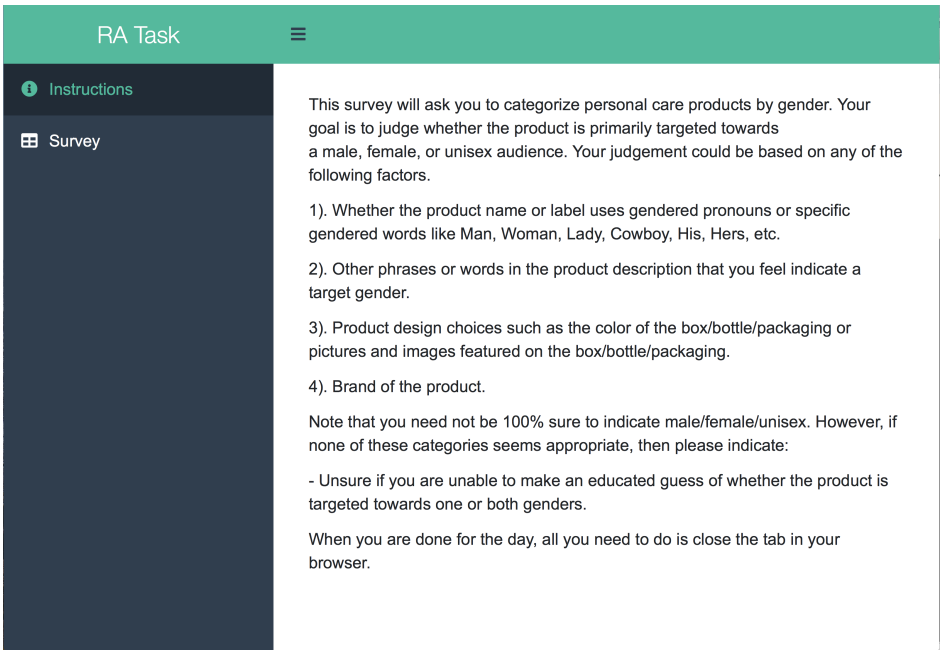
Notes: Screenshots taken on `Walgreens.com` on September 1, 2020.

B.2 Hand-Coding Product Images

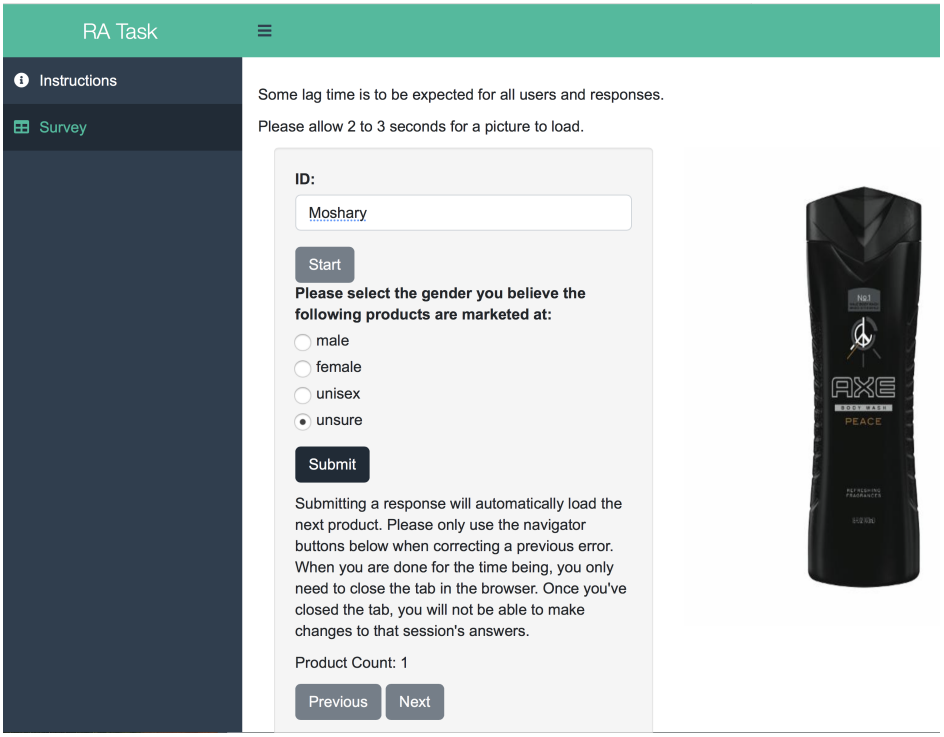
We recruited six undergraduates at The University of Chicago and Northwestern University to assign gender labels to personal care product images from Label Insight. The research assistants were selected based on their performance on a 25 image training dataset, where their answers were compared to our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

Figure A2: Webapp for Gender Classification

(a) Instructions



(b) Task



B.3 Panelist Purchases

The 2006-2018 Nielsen Consumer Panel 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 it 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 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 - Personals - Personal	0.03%
Hair Coloring	1.05%
Hand & Body Lotions	8.83%
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%
Shaving Cream	8.09%

C Product Ingredients

Some of our analyses incorporate information on product ingredients from Syndigo. The Syndigo data does not include all of the products in the Nielsen data, but as Table A2 shows, the products for which we observe ingredients account for the majority of sales in the Nielsen data.

Table A2: Coverage of Ingredients Data

Module	Men	Women
Soap - Bar	90.7%	93.1%
Soap - Specialty	90.4%	91.5%
Deodorants - Personal	96.2%	96.4%
Hair Coloring	75.8%	71.7%
Shampoo	91.0%	74.0%
Shaving Cream	88.8%	77.4%
Overall	88.8%	84.0%

Notes: This table describes the market share of products for which we have ingredients data. Market shares are calculated using product sales in the Nielsen data from 2015-2018.

Our estimates of the pink tax control for the active and top five inactive ingredients. As a robustness check, we also estimate price differences controlling for active and top three inactive ingredients. We view these as conservative interpretations of the Pink Tax Repeal Act's concept of substantial similarity in product materials because most of the products that we study comprise many more than five ingredients. Table 2 reports the median number of inactive ingredients by category, which range from 10 (deodorants) to 55 (hair coloring). Table A3 lists the the most prevalent top five ingredients by category and gender.

Table A3: List of Most Prevalent Top Five Ingredients by Category and Gender

Product Module	Ingred. Type	Rank	Men	Women
Bar Soap	Inactive	1	water	water
Bar Soap	Inactive	2	sodium tallowate	sodium tallowate
Bar Soap	Inactive	3	sodium cocoate	sodium lauryl isethionate
Bar Soap	Inactive	4	glycerin	stearic acid
Bar Soap	Inactive	5	sodium palm kernelate	glycerin
Body Wash	Inactive	1	sodium laureth sulfate	water
Body Wash	Inactive	2	water	sodium laureth sulfate
Body Wash	Inactive	3	fragrance	cocamidopropyl betainee
Body Wash	Inactive	4	cocamidopropyl betainee	fragrance
Body Wash	Inactive	5	sodium chloride	glycerin
Deodorant	Active	1	alum. zirc. tetrachlorohydrex gly	alum. zirc. tetrachlorohydrex gly
Deodorant	Active	2	alum. zirc. trichlorohydrex gly	alum. chlorohydrate
Deodorant	Active	3	alum. chlorohydrate	alum. zirc. trichlorohydrex gly
Deodorant	Active	4	alum. zirc. octachlorohydrex gly	alum. zirc. octachlorohydrex gly
Deodorant	Active	5	alum. zirc. pentachlorohydrex gly	alum. sesquichlorohydrate
Deodorant	Inactive	1	propylene glycol	cyclopentasiloxane
Deodorant	Inactive	2	alcohol	alcohol
Deodorant	Inactive	3	cyclopentasiloxane	ppg-14 butyl ether
Deodorant	Inactive	4	water	dimethicone
Deodorant	Inactive	5	ppg-14 butyl ether	water
Hair Coloring	Inactive	1	water	water
Hair Coloring	Inactive	2	hydrogen peroxide	propylene glycol
Hair Coloring	Inactive	3	alcohol	hydrogen peroxide
Hair Coloring	Inactive	4	ethanolamine	isopropyl alcohol
Hair Coloring	Inactive	5	cetyl alcohol	ethoxydiglycol
Shampoo	Inactive	1	sodium laureth sulfate	sodium laureth sulfate
Shampoo	Inactive	2	water	water
Shampoo	Inactive	3	cocamidopropyl betainee	cocamidopropyl betainee
Shampoo	Inactive	4	sodium chloride	sodium chloride
Shampoo	Inactive	5	ammonium lauryl sulfate	glycol distearate
Shaving Cream	Inactive	1	water	water
Shaving Cream	Inactive	2	triethanolamine	triethanolamine
Shaving Cream	Inactive	3	stearic acid	palmitic acid
Shaving Cream	Inactive	4	palmitic acid	stearic acid
Shaving Cream	Inactive	5	isopentane	isopentane

Notes: This table lists the ingredients that most frequently appear in the top five ingredients for products in each category and gender. It is derived from Syndigo data.

C.1 Robustness to Definition of Product Formulation

Our main results define a product formulation as the combination of manufacturer, active ingredient, and the first five inactive ingredients where the order of ingredients matters. In this section we present results that show the robustness of our main findings to alternative definitions of formulation.

First, we consider the extent of formulation overlap in stores' product assortments using an alternative definition of formulation that relaxes the set of ingredients to the top three. The share of products with overlapping formulations increases under this alternative definition, but the main conclusion that most products do not have a formulaic analog offered to the other gender still holds.

Table A4: Overlap in Manufacturer-Top Three 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	12.4%	20	18.8%	1,040	15.1%
Bar Soap	women	11	10.5%	29	14.8%	1,500	18.1%
Body Wash	men	17	50.0%	44	51.9%	1,658	52.9%
Body Wash	women	40	23.6%	87	36.2%	2,757	45.3%
Deodorants	men	34	41.5%	116	52.8%	4,069	48.7%
Deodorants	women	34	42.5%	103	61.4%	3,714	56.6%
Hair Coloring	men	4	0.0%	15	0.0%	243	0.0%
Hair Coloring	women	15	0.0%	167	0.0%	1,753	0.0%
Shampoo	men	14	29.2%	35	41.2%	961	40.6%
Shampoo	women	46	8.3%	132	28.3%	3,356	30.5%
Shaving Cream	men	12	25.3%	31	31.1%	1,360	27.7%
Shaving Cream	women	4	77.4%	10	90.0%	346	94.7%

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. In this robustness, we consider an alternative definition of a formulation as the combination of manufacturer, active ingredient, and top three 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 ingredient information in the Syndigo data. Convenience stores are dropped because they have very small assortments.

Next, we consider the robustness of our pink gap and tax estimates to an alternative definition of formulation in which the order of the top three or five ingredients does not matter. As shown in Table A5 and Figure A3, this alternative definition of formulation produces very similar results as our original approach where formulation is defined by the order of the top ingredients.

Figure A3: Comparison of Pink Tax Estimates
When Order Does/Does Not Matter

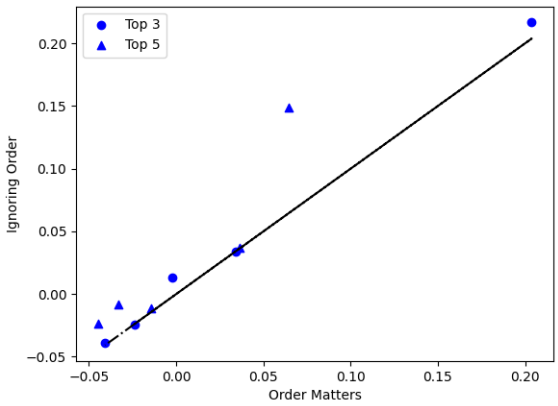


Table A5: Pink Tax by Category and Formulation Definition, 2015-2018

	(1)	(2)	(3)	(4)
	Unit	Unit	Unit	Unit
Module	Shelf Price	Shelf Price	Shelf Price	Shelf Price
Bar Soap	0.03***	0.03***	0.04***	0.04***
	(0.00)	(0.00)	(0.00)	(0.00)
	0.23	0.23	0.23	0.23
	14.9%	14.9%	15.9%	16.1%
	7,273,999	7,273,999	7,273,999	7,273,999
	125	115	157	150
Body Wash	-0.02***	-0.02***	-0.01**	-0.01**
	(0.00)	(0.00)	(0.00)	(0.00)
	0.28	0.28	0.28	0.28
	-8.4%	-8.4%	-5.1%	-4.1%
	18,264,364	18,264,364	18,264,364	18,264,364
	244	223	400	354
Deodorant	0.20**	0.22**	0.06	0.15*
	(0.05)	(0.06)	(0.06)	(0.05)
	1.51	1.51	1.51	1.51
	13.5%	14.4%	4.3%	9.9%
	29,727,058	29,727,058	29,727,058	29,727,058
	266	244	349	320
Shampoo	0.00	0.01	-0.03***	-0.01
	(0.01)	(0.01)	(0.00)	(0.01)
	0.51	0.51	0.51	0.51
	-0.5%	2.6%	-6.5%	-1.7%
	23,296,620	23,296,620	23,296,620	23,296,620
	317	296	575	509
Shaving Cream	-0.04***	-0.04***	-0.04***	-0.02**
	(0.00)	(0.00)	(0.00)	(0.00)
	0.57	0.57	0.57	0.57
	-7.2%	-6.9%	-7.9%	-4.2%
	5,746,248	5,746,248	5,746,248	5,746,248
	132	128	176	166
Data	Syndigo	Syndigo	Syndigo	Syndigo
Formulation FE	Top 3	Top 3	Top 5	Top 5
Order Ingred. Matters	Y	N	Y	N

Notes: This table exclude razors because their ingredients are not reported and hair coloring because there is insufficient overlap in ingredients across men's and women's products. Columns (1) and (2) define a formulation using the top 3 ingredients, while columns (3) and (4) use the top 5. Columns (1) and (3) coincide with the specifications reported in Table 5 of the draft, and incorporate order of ingredients when defining a formulation. Columns (2) and (4) show robustness to defining formulation without specifying the order in which the ingredients occur. For each category, the first row reports the average price difference and the second row reports the standard error (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 difference, calculated as the ratio of row one to row three. The fifth row reports the number of observations. The sixth row reports the number of unique formulations. Regressions are estimated separately by product module and include store and year fixed effects.*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

D Quantity Discounts

This section explores the distribution of package sizes for men's and women's products and reports estimates of the pink tax that control for product size. This analysis sheds light on the extent to which firms differentiate products targeted at men vs women using package size, and it also gives a sense for whether we can plausibly disentangle gender price differences from quantity discounts.

D.1 Descriptives on Product Size

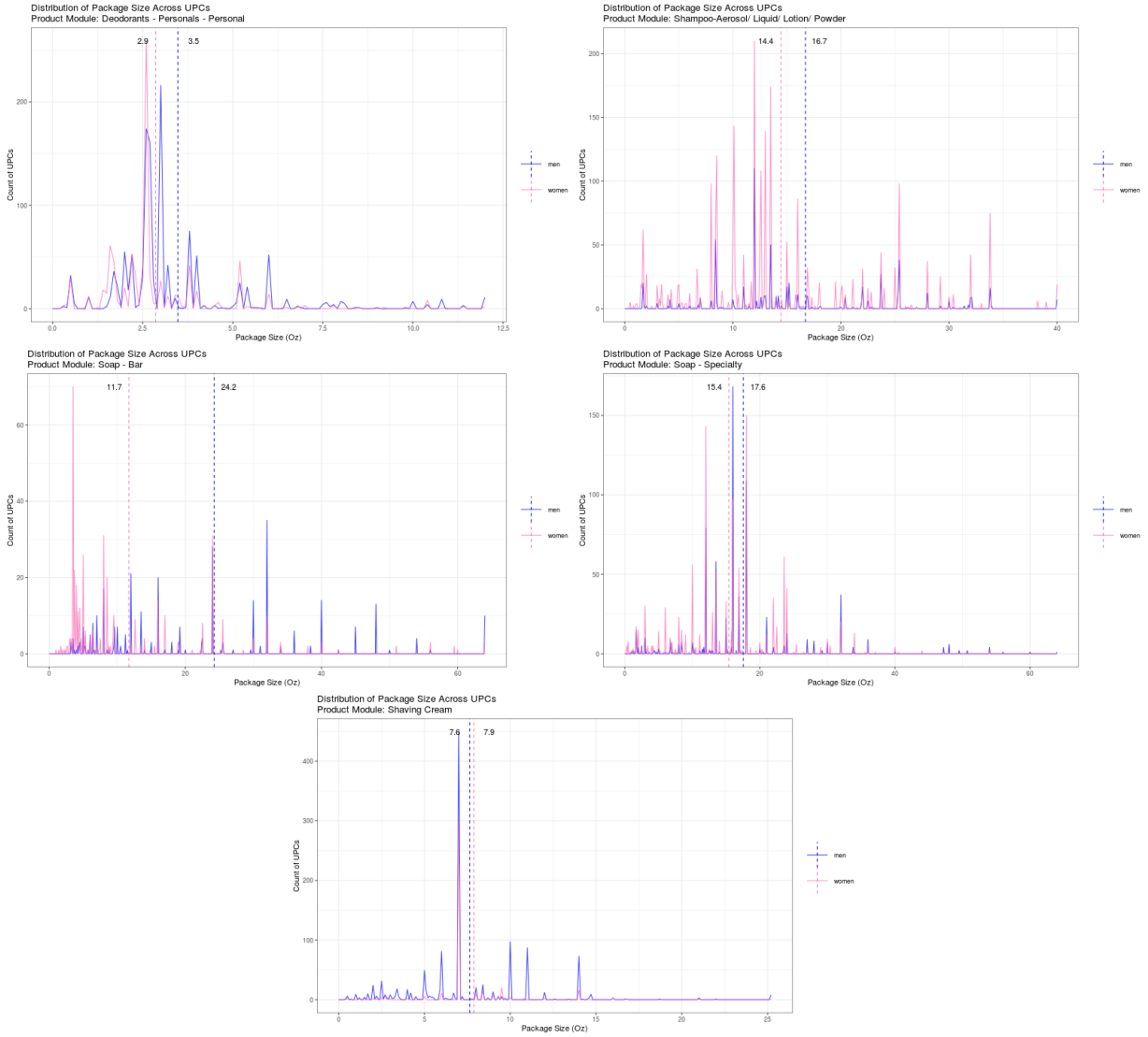
First, we plot the distribution of package size of products targeted to men and women. Figures A4 and A5 show the distribution of package sizes for each category. The vertical dashed lines indicate the mean package size for products targeted to each gender. For all but two categories (shaving cream and non-disposable razors), the products targeted to men are on average larger than the products targeted to women. However, this analysis pools products across manufacturers, including manufacturers that only sell products targeted to one gender. Because our estimates of the pink gap and pink tax rest on within-manufacturer comparisons, we also compute the average within-manufacturer difference in package size for each module. These results are reported in Table A6. Overall, we find that within-manufacturer, men's products tend to be larger. This pattern underscores that size is one way that firms differentiate the products that they target at men and women.

Table A6: Average Within-Manufacturer Difference in Pack-size

Module	Avg Diff. Within Manuf.
Bar Soap	7.80
Body Wash	0.29
Deodorants - Personal	0.23
Hair Coloring	0.02
Razor Blades	0.82
Razors Disposable	1.07
Razors Non-Disposable	-0.09
Shampoo	1.21
Shaving Cream	1.09

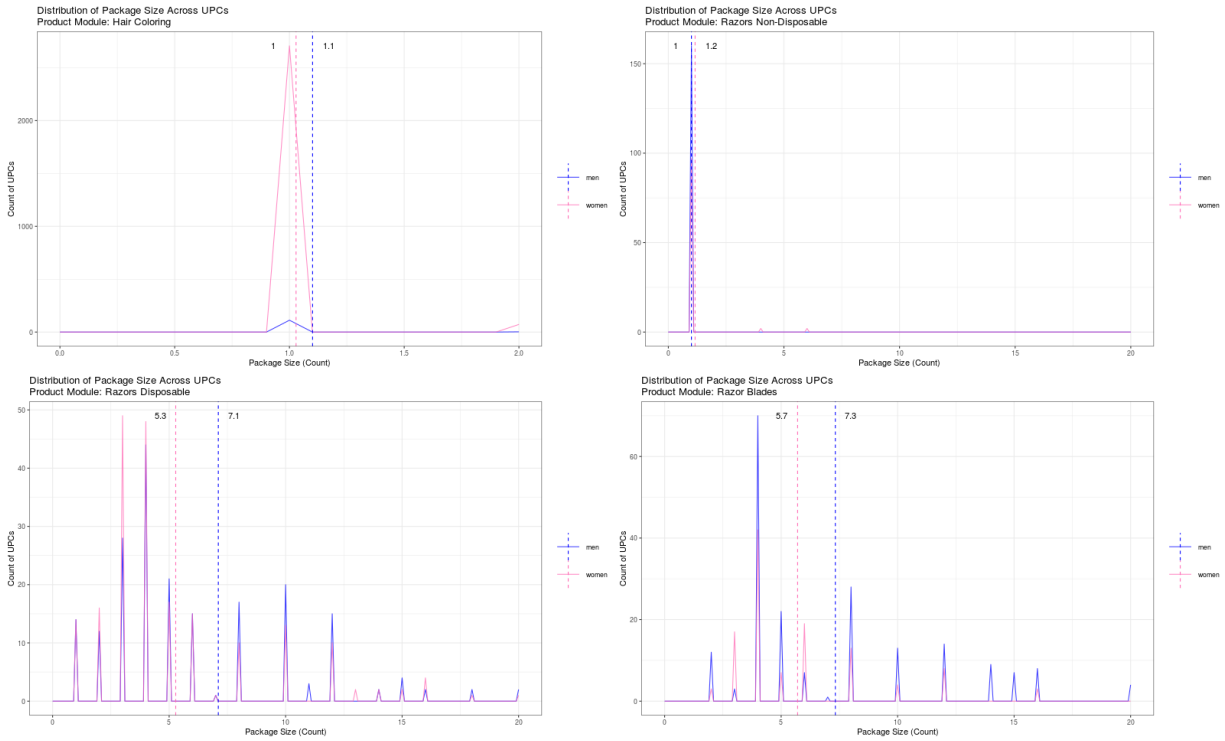
Notes: For each module and manufacturer, we compute the average pack-size of the products that are targeted to each gender and then calculate the difference in these averages, $AvgMen'sSize_{cm} - AvgWomen'sSize_{cm}$. Then within each module, we compute the average difference across manufacturers. Units are in counts for Hair Coloring and Razor products and in ounces for all other product modules.

Figure A4: Distribution of Package Size by Gender Target and Category



Notes: Package size records either the number of ounces contained in each product. The unit of observation is a UPC. Distributions shown separately for products targeted to men and women. The vertical dashed lines and associated labels indicate the mean package size targeted to each gender.

Figure A5: Distribution of Package Size by Gender Target and Category, Continued



Notes: Package size records the count of items contained in each product. The unit of observation is a UPC. Distributions shown separately for products targeted to men and women. The vertical dashed lines and associated labels indicate the mean package size targeted to each gender.

D.2 Price Differentials Controlling for Product Size

Next, we re-estimate the specifications reported in Table 5, adding package size as an independent variable in our main specification (1). Results are shown in Table A7.

As expected, controlling for size tends to shift our pink gap estimates down. This analysis indicates that, as per our intuition, larger products are typically less expensive per ounce. Thus, because men's products are larger on average, when we account for quantity discounts, the pink gap shrinks. In some cases, the sign of our pink gap estimate even flips from positive to negative.

Ultimately, we see the estimates *unconditional* on product size as best aligned with how regulators and consumers conceptualize a pink gap/tax, because these estimates capture the price differential that consumers face for the products targeted to men vs. women. In other words, the analysis above reveals that quantity discounts are economically meaningful for personal care products, but also that these discounts are *differentially offered* for men's products.

Table A7: Price Gap by Category with Size Controls, 2015-2018

	(1)	(2)	(3)	(4)
	Unit	Unit	Unit	Unit
Module	Shelf Price	Shelf Price	Shelf Price	Shelf Price
Bar Soap	0.04***	0.04***	0.00*	0.00**
	(0.00)	(0.00)	(0.00)	(0.00)
	0.23	0.23	0.23	0.23
	15.7%	18.3%	0.6%	2.0%
	7,832,795	7,273,999	7,273,999	7,273,999
Body Wash	0.03**	0.04**	-0.02***	-0.02***
	(0.01)	(0.01)	(0.00)	(0.00)
	0.28	0.28	0.28	0.28
	11.5%	13.2%	-8.6%	-7.4%
	20,076,939	18,264,364	18,264,364	18,264,364
Deodorant	0.28***	0.28***	0.10	-0.02
	(0.01)	(0.01)	(0.05)	(0.06)
	1.50	1.51	1.51	1.51
	18.5%	18.3%	6.8%	-1.4%
	31,001,944	29,727,058	29,727,058	29,727,058
Hair Coloring	0.40			
	(0.18)			
	8.49			
	4.7%			
	28,450,230			
Razor Blades	0.22**			
	(0.04)			
	3.68			
	5.9%			
	3,997,551			
Razors Disposable	-0.32**			
	(0.07)			
	2.28			
	-14.2%			
	6,014,671			
Razors Non-Disposable	-0.26			
	(0.18)			
	11.69			
	-2.3%			
	3,158,478			
Shampoo	-0.02*	-0.07***	-0.01*	-0.03***
	(0.01)	(0.01)	(0.00)	(0.00)
	0.50	0.51	0.51	0.51
	-4.7%	-12.9%	-2.5%	-5.4%
	30,835,605	23,296,620	23,296,620	23,296,620
Shaving Cream	-0.03**	-0.04***	-0.04***	-0.06***
	(0.01)	(0.01)	(0.00)	(0.00)
	0.59	0.57	0.57	0.57
	-5.8%	-7.6%	-7.7%	-10.0%
	6,773,794	5,746,248	5,746,248	5,746,248
Data	All	Syndigo	Syndigo	Syndigo
Manufacturer FE	Y	Y	N	N
Formulation FE	N	N	Top 3	Top 5

Notes: The sample in column (1) comprises the full set of products. Columns (2)-(4) exclude razors because their ingredients are not reported and hair coloring because there is insufficient overlap in ingredients across men's and women's products. Columns (3) and (4) include formulation fixed effects. For each category, the first row reports the average price difference and the second row reports the standard error (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 difference, calculated as the ratio of row one to row three. The fifth row reports the number of observations. Regressions are estimated separately by product module and include product size measured in ounces (counts for razors), as well as store and year fixed effects.*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

E Analyses of Household-Level Purchase Data

E.1 Cross-Segment Purchase Behavior

This section documents the extent to which households in the HMS data that are comprised of a single man or single woman buy products that are targeted to their own gender. For each product module, we restrict to the set of households that purchased at least one product in that module. Table A8 presents the results. Panel (1) reports the average number of times that households buy any product in that module separately by gender target. For example, the table indicates that single men that buy deodorant on average purchase a deodorant targeted to men 2.81 times over the course of a year and buy a deodorant targeted to women 0.53 times a year. The pattern is flipped for single women, who on average buy a deodorant targeted to women 2.44 times and a deodorant targeted to men 0.65 times a year. Panel (2) reports the number of unique products (UPCs) that are purchased. On average, single men and women buy about 1.65 unique UPCs targeted to their own gender, so there is some repeat purchasing as well as some substitution within products targeted to one's own gender. Across modules, most consumers primarily purchase products that are targeted to their own gender. Exceptions include men's purchases of bar soap, hair coloring, and shampoo, where we find that they buy products targeted to women almost as often as they buy products targeted to men. For women, the exception is shaving cream, where women are more likely to buy a product targeted to men than to women. Overall, the results suggest that gender targeting is quite effective, though it does not perfectly segment the market.

Table A8: **Own and Cross Gender Purchasing Behavior by Module**

		(1)		(2)	
		Number of Purchases		Number of Unique Products	
Module	Household Type	Men's	Women's	Men's	Women's
Bar Soap	Single Man	1.74	1.00	1.11	0.51
Bar Soap	Single Woman	0.83	1.82	0.56	1.07
Body Wash	Single Man	2.56	1.39	1.53	0.81
Body Wash	Single Woman	0.45	2.99	0.28	1.97
Deodorants	Single Man	2.81	0.53	1.69	0.31
Deodorants	Single Woman	0.65	2.44	0.41	1.59
Hair Coloring	Single Man	2.24	2.35	0.84	1.06
Hair Coloring	Single Woman	0.03	4.69	0.02	2.33
Razor Blades	Single Man	1.65	0.05	1.20	0.04
Razor Blades	Single Woman	0.53	0.97	0.40	0.77
Razors Disposable	Single Man	2.25	0.41	1.39	0.21
Razors Disposable	Single Woman	0.65	1.59	0.40	1.11
Razors Non-Disposable	Single Man	1.33	0.28	1.15	0.12
Razors Non-Disposable	Single Woman	0.47	1.05	0.35	0.86
Shampoo	Single Man	0.98	1.56	0.70	0.95
Shampoo	Single Woman	0.23	2.48	0.17	1.75
Shaving Cream	Single Man	2.31	0.07	1.47	0.04
Shaving Cream	Single Woman	1.33	0.81	0.90	0.60

Notes: Annual purchase metrics reported using data from 2018 for single male and female households in the Nielsen HMS data. Analyses for each module include the set of households that made at least one purchase in that module.

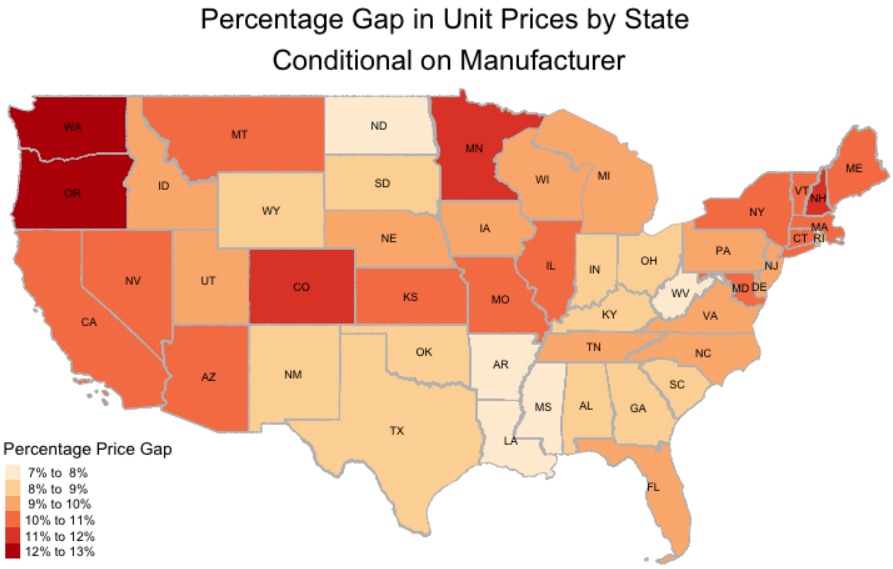
E.2 Calculation of the Gains from Switching

To approximate household savings from switching to a cheaper product targeted to the other gender, we first compute the dollar spending, average price, and total volume (measured in ounces or counts) of purchases made by each household for each product category/gender combination analyzed in Table 5. 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 5. When estimating savings from switching to a comparable formulation, we use the estimates in column (4), and when estimating savings when switching within manufacturer across formulations, we use the estimates in column (1). 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 4). The estimated savings from switching within formulation across gender ($<$1%) are much lower than the potential savings from switching across formulations (9%) both because most purchases do not 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.

F Heterogeneity across States

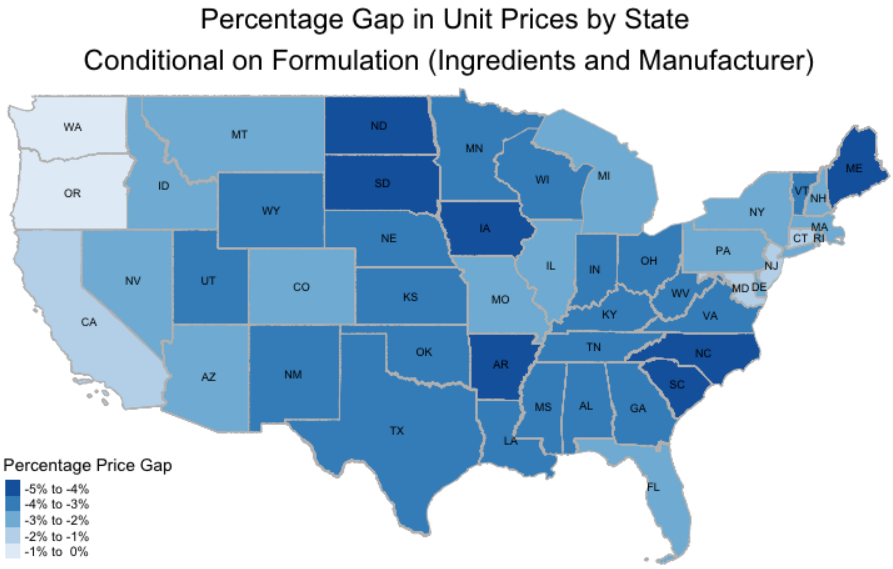
This appendix explores whether and to what extent gender price differences vary across the country. Figure A6 maps the gender price gap separately by state for the contiguous US. The estimates are based on a pooled regression of log prices on an indicator for whether a product is targeted at women. The regression includes manufacturer$\times$category, store$\times$category, and year$\times$category fixed effects. Percent price differences are calculated as $\exp(\hat{\beta})-1$, where β is the coefficient on the gender targeting indicator. The unconditional gender price difference is large and positive across the board, ranging from 7% to 13%. Figure A7 presents state-level estimates of the pink tax, which control for formulation (defined as the combination of manufacturer, active, and top five inactive ingredients) fixed effects. As for the national estimate, the pink tax is negative and economically small for all states.

Figure A6: Pink Gap Estimates by State



Notes: This map shows our estimate of the unconditional gender price gap for each state for the 2015-2018 period. Estimates are recovered from product-store-year level regressions of log price on an indicator for whether the product is targeted at women. Controls include manufacturer×category, store×category, and year×category fixed effects.

Figure A7: Pink Tax Estimates by State



Notes: This map shows our estimate of the pink tax for each state for 2015-2018. Estimates are recovered from product-store-year level regressions of log price on an indicator for whether the product is targeted at women. Controls include fixed effects for formulation (defined as the combination of manufacturer, active, and top five inactive ingredients), store×category and year×category.

G NYC DCA Report Replication and Extension

In this section, we revisit evidence from Bessendorf (2015), a NYC Department of Consumer Affairs (NYC DCA) 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 first replicate the results of the report using the original data collected by the NYC DCA for the study. This data was collected for 61 pairs of men's and women's UPCs sold in NYC drugstores in 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 A9: 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	($)	($)	Tax	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%		
Lotion	Drugstores	NYC Only	-0.07**	7.80	-0.9%	11.0%	-8.0%
	Drugstores	National	-0.18***	7.46	-2.4%		
	All	National	-0.09***	6.55	-1.4%		
Razor	Drugstores	NYC Only	0.78***	10.60	7.4%	9.3%	12.3%
	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%	11.3%
	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. The prices of these matched products in the Nielsen data comprise the sample in columns (3)-(5).

Table A9 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 rescales prices using the ratio of sizes.[31] It arrives at a 13% pink tax via a simple average across categories. Column

[31] The report does not rescale prices for body wash. Because our aim is to replicate their methodology, estimates in Table A9

(6) replicates Bessendorf (2015)'s estimates using the original data collected by the NYC DCA.[32] Based on the NYC DCA data, women's products are more expensive in six out of seven 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 excludes 19 of the products in the NYC DCA sample (10 of them are private label products) because we cannot match both products in the pair to a product in the Nielsen data. We do not believe this substantively affects our estimates of price differences; column (7) shows that the matched UPCs produce similar estimates of price differences in the NYC DCA data in all but one product category, lotions. Column (3) reports average price difference in dollars for different samples, and column (5) reports the implied pink tax. The estimates at the national level echo Bessendorf (2015) in that five of the seven 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. Table A10 reports the market share of these products in the 2015 Nielsen RMS data by category. As shown in column (1), across all categories, the share is modest, ranging from 2.3% of shampoo 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 that 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. Thus, even if the individual products included in the sample were representative of their respective brands' pricing strategies, a large share of the market is omitted.

Table A10: Market Share of UPCs Studied in the NYC DCA Report

	(1)	(2)
Category	UPCs Market Share	Brands Market Share
Bodywash	4.1%	32.2%
Deodorant	5.3%	35.4%
Lotion	3.2%	16.7%
Razors	12.4%	23.0%
Shampoo	2.3%	19.3%
Shaving Cream	19.7%	48.8%
Total	5.5%	28.6%

A second concern is that the products studied in the NYC DCA report were not selected at random. 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

do not rescale in this category either.

[32] We replicate all values in Bessendorf (2015) except the average price of razors targeted to women. Bessendorf (2015) reports an average price of $8.90 for women's razors, while we find an average price of $8.73. We believe the discrepancy is likely due to a typo in the product-level price data or a mistake in computing the averages in Bessendorf (2015).

Table A11: 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	3	2	1
Lotion	2	0	-	2	2	0	0	0
Shaving Cream	6	0	-	6	4	4	1	0
Total	32	11	91%	97%	78%	62%	44%	28%

Notes: Column (1) reports the number of product pairs that we could identify in the Nielsen and Syndigo datasets. 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.

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 NYC DCA 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.[33] Table A11 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 (relevant only in certain categories), column (4) reports the number with the same active and first inactive ingredients, etc.[34] Less than one-third of product pairs comprise the same top 5 ingredients. 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 A9 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 whether the 61 product pairs considered are representative. We provide more details of our analysis of the NYC DCA report below for interested readers.

Identifying the UPCs of Products in the NYC DCA Report

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.

[33] Only one product pair identified in the Nielsen data does not have Syndigo ingredient information.

[34] The FDA requires active ingredients be reported first, then inactive ingredients in order of predominance. Any order is permitted for inactive ingredients comprising less than 1% of the product. [https://www.fda.gov/cosmetics/cosmetics-labeling-regulations/cosmetics-labeling-guide#clgl]

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.
3. For the 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.

Additional Notes on Estimating Price Differences

We also follow 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 rescale 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.