Web Appendix

1 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.

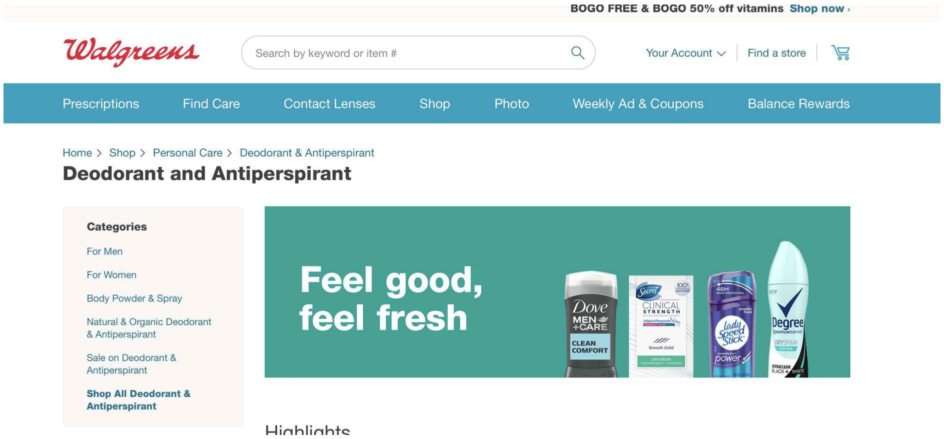
2 Gender-Targeting Data Sources

2.1 Walgreens

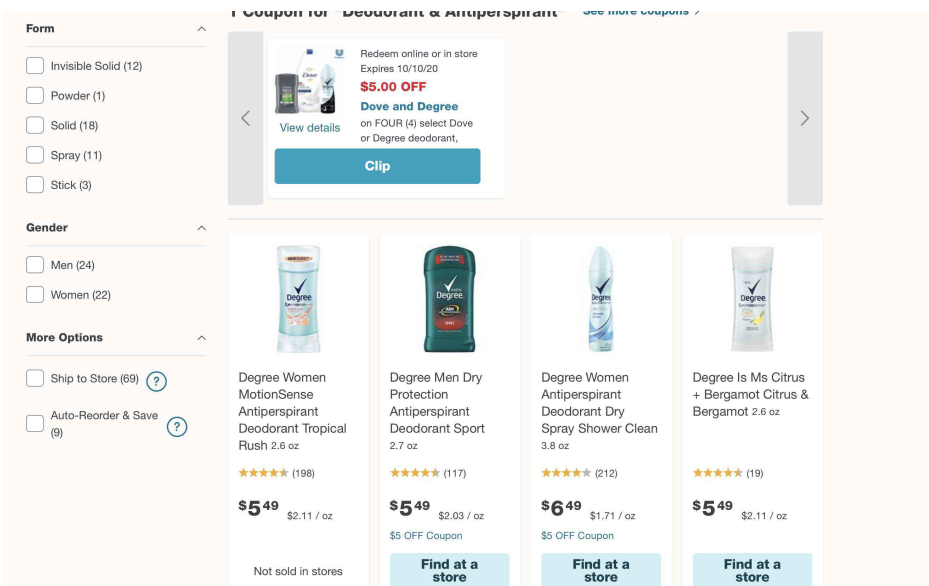
We extract gender information from the Walgreens website. The website explicitly categorizes certain product categories by gender. Figure A1 (a) presents one such example for the Deodorant & Antiperspirant category. We also collect gender information from search result page gender filters, as in Figure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

Figure A1: Walgreens.com Gender Categorizations

(a) Primary Gender Classification



(b) Gender Filter on Search Results Page



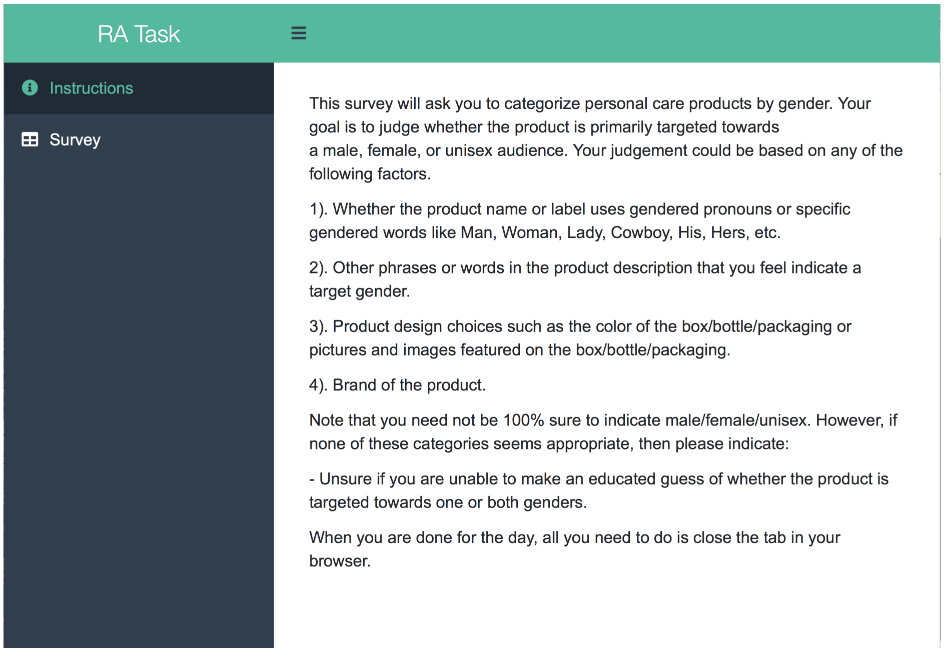
Notes: Screenshots taken on `Walgreens.com` on September 1, 2020.

2.2 Hand-Coding Product Images

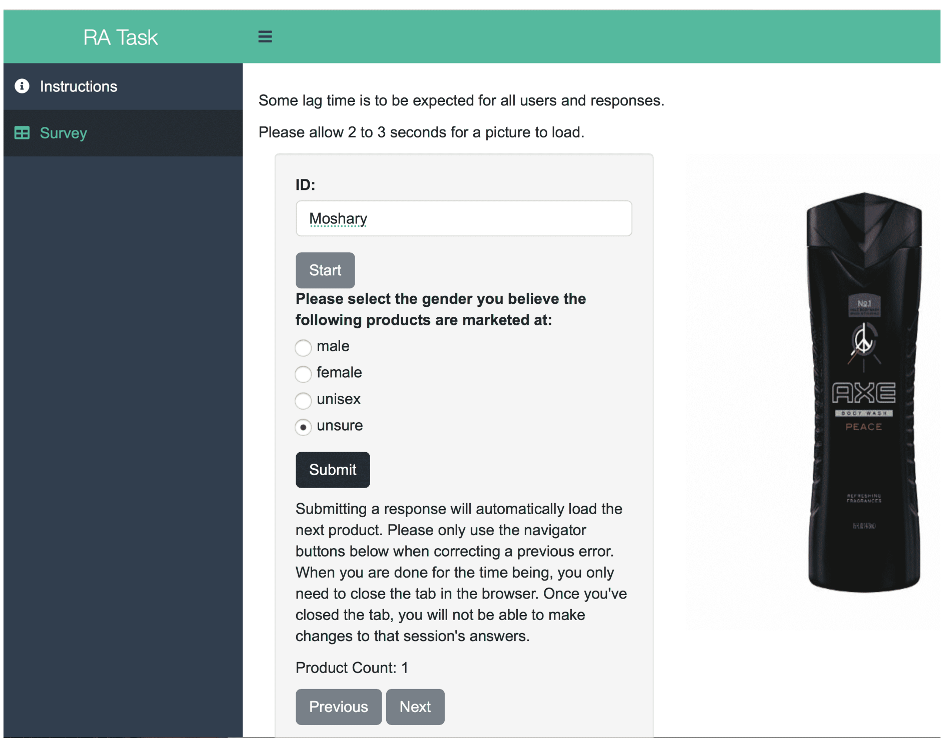
We recruited six undergraduates at The University of Chicago and Northwestern University to assign gender labels to personal care product images from Label Insight. The research assistants were selected based on their performance on a 25 image training dataset, where their answers were compared to our 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



2.3 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.

2.4 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 2.3. 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%

3 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.

3.1 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

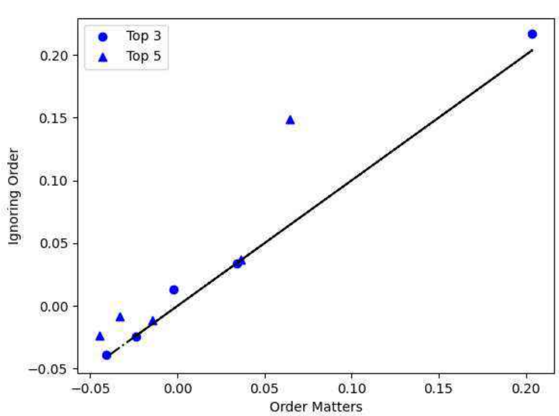


Table A5: Pink Tax by Category and Formulation Definition, 2015-2018

Module	(1) Unit Shelf Price	(2) Unit Shelf Price	(3) Unit Shelf Price	(4) Unit 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.*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

4 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.

4.1 Descriptives on Product Size

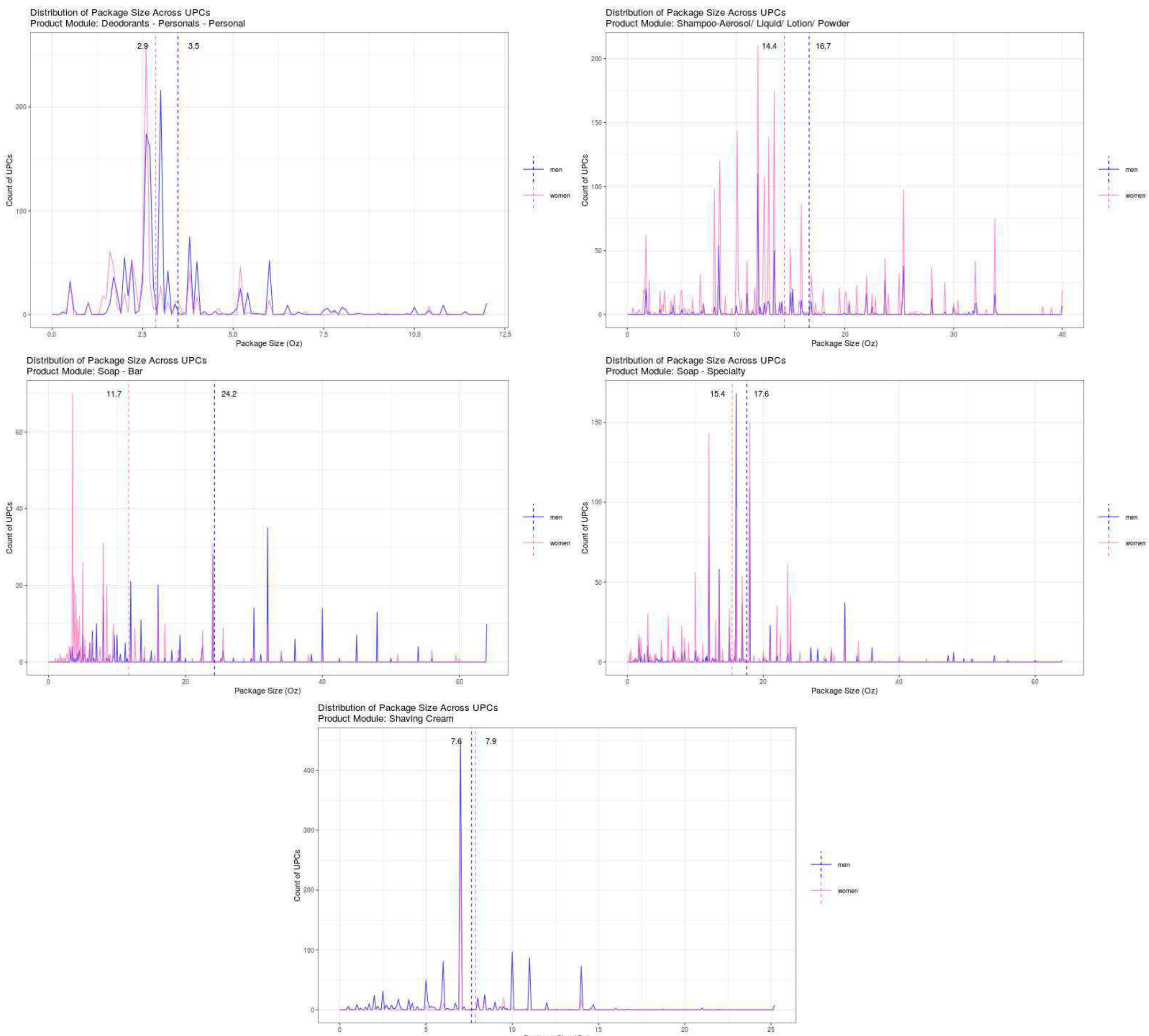
First, we plot the distribution of package size of products targeted to men and women. Figures A4 and A5 show the distribution of package sizes for each category. The vertical dashed lines indicate the mean package size for products targeted to each gender. For all but two categories (shaving cream and 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

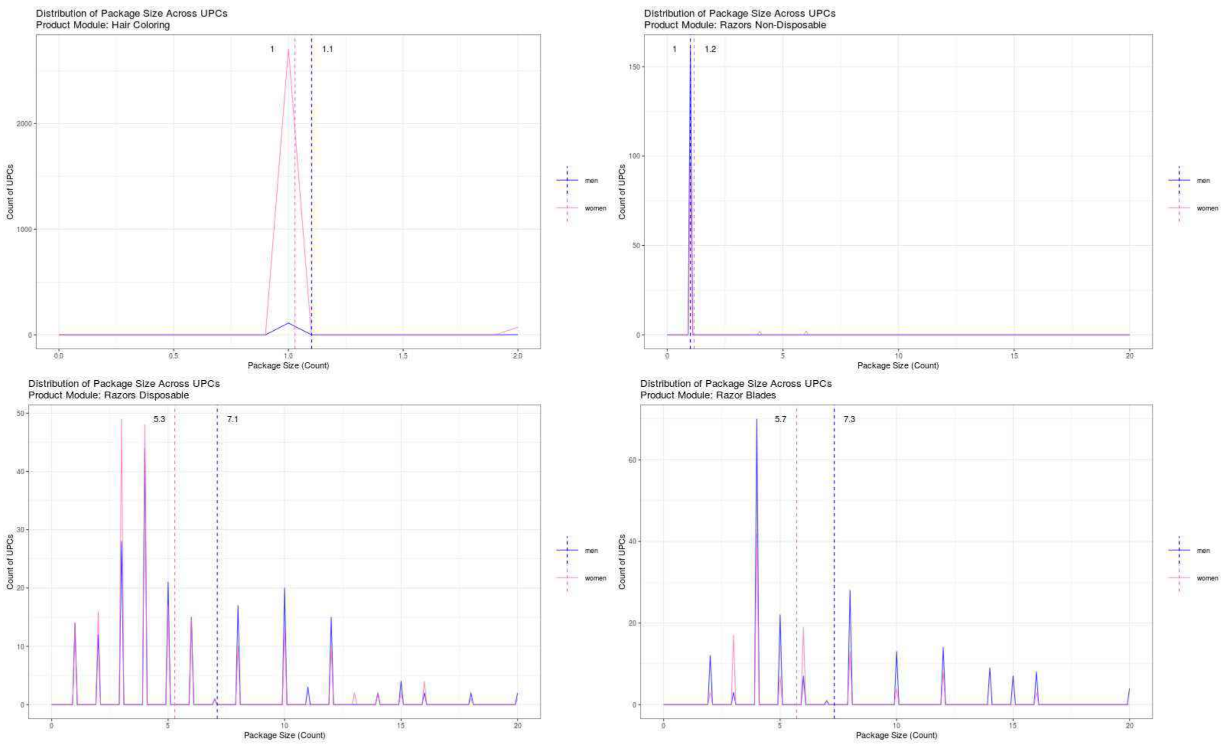
Notes: For each module and manufacturer, we compute the average pack-size of the products that are targeted to each gender and then calculate the difference in these averages, $AvgMen'sSize_{cm} - AvgWomen'sSize_{cm}$. Then within each module, we compute the average difference across manufacturers. Units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

4.2 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

Module	(1) Unit Shelf Price	(2) Unit Shelf Price	(3) Unit Shelf Price	(4) Unit 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.*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

5 Analyses of Household-Level Purchase Data

5.1 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) Number of Purchases		(2) 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.

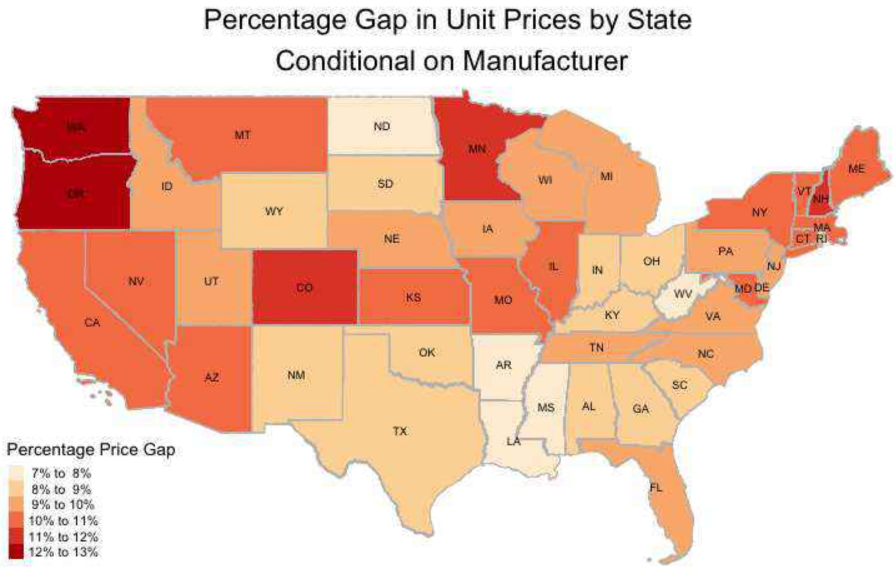
5.2 Calculation of the Gains from Switching

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.

6 Heterogeneity across States

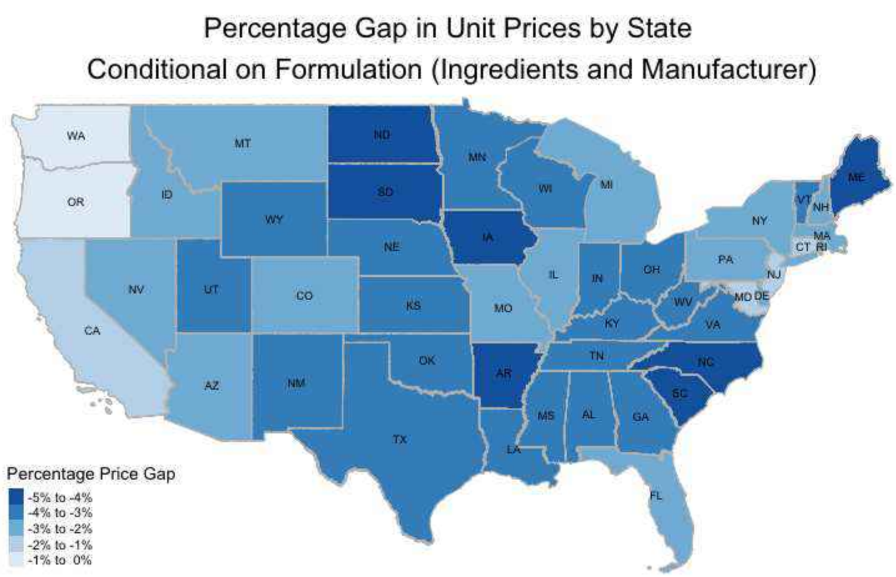
This appendix explores whether and to what extent gender price differences vary across the country. Figure A6 maps the gender price gap separately by state for the contiguous US. The estimates are based on a pooled regression of log prices on an indicator for whether a product is targeted at women. The regression includes manufacturer×category, store×category, and year×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.

7 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.[1] It arrives at a 13% pink tax via a simple average across categories. Column (6) replicates Bessendorf (2015)'s estimates using the original data collected by the NYC DCA.[2] 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

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

[2] 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).

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

Category	(1) UPCs Market Share	(2) 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 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.[3] 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.[4] 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

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

[4]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]

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.

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. 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.

References

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