

# Quantifying Gender Representation in TV Advertising

Clément Bellet

Erasmus School of Economics

Martina Pocchiari

NUS Business School, National University of  
Singapore

This version: June 2024

## Abstract

Advertising plays a crucial role in shaping gender norms, with the potential to either promote inclusivity or reinforce gender stereotypes. This paper examines gender representation in TV advertising using comprehensive unstructured text data describing all TV advertising content in the United States since 2010. We quantify two dimensions of gender representation: as role models (i.e. which gender is represented), and in context (i.e. how gender is represented). We rely on a dual approach: a data-driven method that involves pattern matching and context scoring, and a survey-based method that elicits human associations. We find that gender representation in TV ads has evolved over the past 10 years, with varying degrees of adherence to traditional gender norms across different brands and product categories. We analyze the extent to which advertisers rely on gender-congruent contexts, and align gender representation in their ads with the identities of their target audiences. Finally, we highlight instances where gender representations diverge from audience identities and discuss potential explanations for these discrepancies.

**Keywords:** TV Advertising, Gender Representation, Text-as-Data, Visual Marketing

---

The authors thank Raphael Barthès for his excellent research support, and seminar participants at EMAC 2024 Annual Conference. The analyses in this paper are calculated (or derived) based in on data from Nielsen Consumer LLC and marketing databases provided through the Nielsen Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from these data only reflect the work of the researchers, and do not reflect the views of Nielsen nor the brands mentioned in the dataset. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein. Corresponding author: Clément S. Bellet, Erasmus School of Economics, Burgemeester Oudlaan 50, 3062 PA Rotterdam. Email: <[bellet@ese.eur.nl](mailto:bellet@ese.eur.nl)>.

# 1 Introduction

Mass media and TV advertising play an important role in promoting positive, inclusive, and diverse gender norms (UNICEF and UN Women 2023). On the one hand, gender representation in TV ads can be a powerful tool for brands and advertisers to promote products, connect with their consumers and audience, communicate brand values, and advance diversity, equity, and inclusion commitments (DEI, Kim et al. 2023). On the other hand, TV ads may depict people in roles and contexts that reflect gender stereotypes, perpetuate biased representations of gender roles in society, and discourage consumers from engaging or buying from the advertiser. For instance, while a TV ad may represent a woman or girl as a role model, their portrayal in a highly stereotypical context may not support the promotion of positive gender norms.

Discriminatory gender representations in advertising can lead to dramatic societal consequences, such as the normalization of violence against women and girls through the portrayal of women in unhealthy relationships (UNICEF and UN Women 2023). This recently led the UK’s Advertising Standards Authority to enforce a “gender stereotypes” rule preventing gender representations likely to cause harm or serious or widespread offense.<sup>1</sup> While several ads are banned every year as a result of this rule (Daniel 2024), there is a general lack of systematic empirical evidence on the extent to which advertisers resort to traditional gender representations, and why they do so.

In this paper, we quantify gender representation in TV ads using unstructured text data, and propose a characterization of gender representation that takes into account both *whether* a gender is represented (*gender role model*), and in which *context* the gender is represented (*gender in context*). For this, we rely on datasets widely used in the study of advertising and its effects: the Nielsen Ad Intel and Consumer Panel datasets. From the Nielsen Ad Intel dataset, we use the text transcribing and describing the advertising creative shown in the TV ad. The variable, created by annotators watching the TV ads, describes the content of each scene of the ad in sequence. The annotators consistently report the presence of a person or people in the ad (as well as their perceived gender identity), plus any other images, visuals, concepts, and texts that appear in the scenes. As suggested by Ash et al. (2021), content labeling by human annotators (versus computer vision) is more relevant for the study of diversity representation, as it captures the identity a viewer is likely to assign to a person in an ad.<sup>2</sup>

We propose two complementary approaches to quantifying the representation of gender role

---

<sup>1</sup>The American Marketing Association (AMA) Statement of Ethics also commits to “never stereotype anyone or depict any group (e.g., gender, race, sexual orientation, religious beliefs, etc.) in a negative or dehumanizing way.” (<https://www.ama.org/ama-statement-of-ethics/>)

<sup>2</sup>Text-based methods are also reportedly effective in quantifying gender representation due to the presence of clearly gendered words and concepts (Ash et al. 2021).

models and contexts. The first, a *data-driven* approach, relies entirely on the text descriptions to measure both the presence of role models and the gender context. In particular, using pattern-matching, we detect the presence of a *gender role model* in the TV ad description if the text includes one or more gendered keywords, such as “man, woman, boy, girl” or derivatives of these words, such as “fireman, policeman, schoolgirl”.<sup>3</sup> Then, we quantify gender contexts by assigning a score to every other ungendered word in the ad description, based on whether the word appeared in the same ad as a male or female role model. This score is assigned to each word across all ads. Averaging all the scores assigned to each word within an ad, we obtain a *gender in context* score. This data-driven context score indicates the extent to which an ad uses words and visuals that are likely to co-occur with the presence of male or female role models.

The second approach elicits *human associations* between the ungendered context words in the ads and the gender role model. Specifically, we ask survey respondents to indicate to what extent they associate the (ungendered) ad content with men, women, both, or neither. While the data-driven approach requires that at least some of the ads directly represent a man or a woman to calculate the context score, this approach does not require the presence of a role model to obtain an association score. At the end of this scoring procedure, each TV ad has an indicator for the presence of gender role models and two scores characterizing the rest of the ad context – a data-driven and a survey-based score.

We find that women’s representation as role models on TV increased between 2010 and 2020, while men’s representation as role models mostly stagnated. However, women have been represented as role models more frequently than men in only 8 out of 29 industries, including toiletries and cosmetics, soaps, cleansers, and polishes, drugs and remedies, and household upkeep. In most cases, these industries are related to traditional gender roles identified by the International Social Survey Programme (ISSP 2016). Men have appeared as role models more frequently than women in 21 industries. Some of these industries are also related to traditional gender roles for men, such as business products, computers, office equipment, and stationery (ISSP 2016). Other industries in which men are over-represented compared to women include sporting goods, beer and wine, automobiles and auto accessories, and apparel.

Regarding gender in context, we find that TV ads have increasingly included words and concepts associated with women over time. The increase in female contexts appears especially pronounced within brands. Male and female role models have generally been represented in gender-congruent contexts over the past decade. However, since 2018, TV ad creatives have less frequently placed female (male) role models in contexts that viewers might traditionally

---

<sup>3</sup>For simplicity and tractability, we focus on the representation of the two most common biological genders – male and female.

associate with female (male) role models.

When considering context representation across industries, we find that in toiletries and cosmetics, household upkeep, retail, and pet supplies ads featuring a female role model, the surrounding context is much more likely to include words and concepts associated with women. Conversely, in sporting goods, toys and games, and beer and wine ads featuring a male role model, the surrounding context tends to include words and concepts associated with men. As discussed earlier, advertisers appear to represent people of different genders in contexts that align with traditional gender role models (ISSP 2016). In other industries, the ad context is not always gender-congruent. For instance, in soaps, cleansers and polishes, and household upkeep ads, men are often depicted in more female contexts, even when they serve as role models. Conversely, in automotive, auto accessories, industrial materials, and computer ads, women are represented in more male contexts, even when they appear as role models.

One question is whether gender representation — especially when congruent with traditional gender roles as documented above — is used by advertisers to connect with the consumers most likely to watch the ads or those most likely to buy the advertised product. Our results at the product level suggest some degree of connection between gender-associated contexts and the identity of consumers in intended target audiences. However, given a target audience, we also document several important incongruities and find significant variations in context scores that cannot be simply explained by gender-based targeting. For instance, deodorant, luggage, shotgun, or automobile lubricant ads are watched by a similar share of male and female audiences. However, deodorant and luggage ads systematically exhibit more female contexts, while shotguns and chainsaw ads systematically exhibit more male contexts.

Given a target audience, the residual differences in gender representation across products may simply capture differences in product purchase by gender. For instance, bras or lipsticks are almost exclusively used by women, while golf clubs are mostly used by men. In these cases, we find results consistent with advertisers attempting to connect with their consumer base through gender-congruent representations. However, similar to the case of gender-based targeting, we notice deviations from this pattern. For instance, men spend about 25% less on horticulture and farming products than women, but horticulture and farming product ads systematically display more male contexts and role models.

We also find that a significant extent of gender representation cannot be explained either by the identity of target audiences or by the gender gaps in consumption. For instance, ads for hygiene products, baby products, housewares, and laundry supplies use more female role models and contexts beyond what gender-based targeting and gender gaps in consumption could

explain. Similarly, ads for horticulture and farming, books, red meat, and tools exhibit more male role models and contexts, and this difference cannot be fully explained by targeting and consumption gaps.

We propose several potential explanations for such discrepancies between gender representations and historical viewership and spending. First, in the absence of data-driven insights, advertisers in certain sectors may underestimate or not fully internalize demographic and consumption shifts. For example, in the horticulture and farming products category, we observe a shift toward increased female spending over time, yet the advertising remains predominantly male-associated.

Second, systematic deviations may suggest brands' repositioning efforts, possibly to reach and attract new customer segments. For instance, men and women spent approximately the same amount on deodorants between 2010 and 2020, with no significant trend. However, over time, brands shifted toward more female role models and contexts. Brand positioning attempts may appear more clearly, on average, within neutral product categories. For example, within the liquor category, brands adopt a mix of female (e.g., Baileys), neutral (e.g., Tuaca), or male (e.g., Jägermeister) contexts.

Finally, other forms of systematic deviations may be consistent with perpetuated clichés and stereotypical beliefs about gender roles. For instance, disposable diapers show a progressive closure of the gender consumption gap (from 17% in 2010 to 8% in 2020). However, female representation in this product category consistently increased between 2012 and 2016, with some reversal occurring only between 2017 and 2020. All brands within the “diapers” product category exhibit mostly female contexts, in contrast with a product category like “grills”, where all brands exhibit exclusively male contexts.

These findings contribute to the studies on the representation of social categories and social minorities in marketing and economics (e.g., Aneja et al. 2023, 2024, Ash et al. 2021, Babar et al. 2023, Balakrishnan et al. 2024, Goli and Mummalaneni 2023, Hartmann et al. 2023, Khan and Kalra 2022, Kim et al. 2023, Kirgios et al. 2022, Overgoor et al. 2023, Wang et al. 2022, Bellet et al. 2024). Our study introduces a novel, two-fold approach to quantifying gender representation in TV ads that systematically characterizes both the gender representation and the context in which this representation occurs. Kim et al. (2023) recently measured visual elements and audio messages in TV advertising as confounding factors in estimating the effectiveness of TV advertising with racial minority representation. These elements pertain to the context in which minorities appear, but they were not characterized for their systematic relation to the representation of minorities. Moreover, our approach relies solely on relatively simple text

annotations and does not require computer vision.

The findings offer insights into advertising strategies and policymaking efforts. For advertisers, the study provides a comprehensive overview of the creative tropes used for gender representation and outlines ad creative solutions where the interaction between gender role models and context might be stereotypical. Furthermore, we inform advertisers about cases where they might be systematically misrepresenting their target viewers or consumers, potentially propagating stereotypical views of gender norms in society.

Policymakers are also actively involved in regulating gender representation in marketing communications, especially when gender representation hinges on stereotypical beliefs (Authority 2022). These institutions can already rely on existing research on gender representation in terms of gender role models. However, there is still a lack of evidence regarding the context in which gender role models are represented. This study provides policymakers with a comprehensive account of how advertisers have represented genders in *contexts* over the last decade, and to what extent those contexts might reflect widespread stereotypical beliefs about gender roles in society.

The rest of the paper is structured as follows. In Section 2 we describe the data and the empirical context of our research and detail the measures of the key constructs. In Section 3 we illustrate insights about gender representation of role models and genders in context over 10 years of TV advertising. Section 4 discusses cases of deviation from the target audience or consumer base, and potential explanations. Section 5 concludes.

## 2 Data and Methods

The main sources of data for this study are the Nielsen Ad Intel database and the Nielsen Home-scan household panel data. We complement these databases with a primary data collection.

### 2.1 Nielsen Ad Intel

The Ad Intel database includes product-level TV advertising data for ads aired in 2010–2020 and records the placement of product-level ads for a specific brand on a given channel, market, day, and time. The most important dimensions of the dataset are the creative ID and creative description, the product ID and category, the brand marketing the product, the industry in which the brand operates, and information about the audience of the TV shows (or estimated viewership) in which ads are displayed. In total, we use data on 2,116,760 ad creatives and their descriptions, 1,910 products marketed by 272,738 brands in 112 product categories and 32 industries. Table 1 provides descriptive statistics about the data.

**Table 1:** Descriptive Statistics.

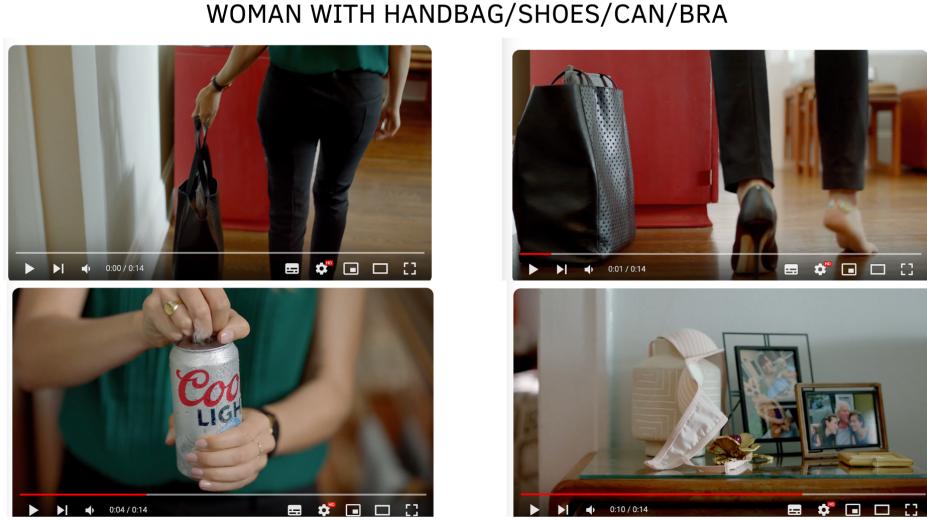
Industry	N. Ads	N. Brands	N. Products	% Male Audience	% Ads
Auto,auto accessories & equipment	618,073	28,306	52	41.699	29.412
Retail	544,634	59,375	313	40.153	25.917
Business & consumer svcs	406,011	71,181	190	40.711	19.321
Drugs & remedies	114,353	24,617	157	39.531	5.442
Entertainment & amusements	114,046	21,917	50	41.593	5.427
Travel,hotels & resorts	71,111	7,735	28	41.234	3.384
Insurance & real estate	46,646	9,195	23	41.135	2.220
Publishing & media	36,048	5,361	43	41.346	1.715
Direct response products	24,378	7,329	10	41.103	1.160
Sporting goods, toys, & games	18,623	8,568	113	47.795	0.886
Foods & food products	18,517	5,187	196	38.616	0.881
Building Material Equipment & Fixtures	13,161	2,621	110	41.303	0.626
Toiletries & cosmetics	12,428	3,642	114	37.162	0.591
Horticulture & farming	9,761	2,637	44	43.372	0.464
Confect., snacks & soft drinks	7,462	1,265	25	41.488	0.355
Beer & wine	6,939	1,141	22	47.437	0.330
Apparel,footwear, & accessories	5,635	1,752	46	46.458	0.268
Household equipment & supplies	5,631	1,542	108	38.757	0.268
Business & Employment Recruit	5,322	2,543	2	41.665	0.253
Computers, office equipment & Stationery	4,201	1,223	45	44.446	0.200
Household furnishing Supplies & Materials	3,234	613	50	39.923	0.154
Gasoline, lubricants & fuels	3,033	533	11	45.322	0.144
Soaps, cleansers & polishes	2,918	672	39	36.500	0.139
Jewelry, optical goods & Cameras	2,679	741	24	44.496	0.127
Pets, pet foods, suppl & organ.	2,457	700	23	39.069	0.117
Industrial materials	1,860	629	10	42.873	0.089
Electrical Entertainment Equipment & Suppl.	1,516	430	31	46.680	0.072
Freight, industrial & Agricultural Dev.	557	115	8	45.966	0.027
Airplanes, aviation services & Equip	118	36	4	46.188	0.006
Misc not elsewhere classified	45	21	10	43.333	0.002
Cigarettes, tobacco & access	22	16	7	41.510	0.001
Tv programs	1	1	1	39.779	0.000

We use the “creative description” variable in the Ad Intel dataset to quantify and characterize gender representation in the ads. The creative description is a brief transcript of the ad content, scene by scene, labeled by Nielsen’s human annotators as they watched the ads. The creative descriptions labeled by human annotators (versus a computer vision model) effectively report the *perceived* identity of the people represented in the ad, as well as their perception of any other concept, visuals, and text that appear in the ad scenes. As reported in Ash et al. (2021), content labeled by human annotators (versus computer vision) is more relevant for the study of representations of gender diversity because it captures the identity a viewer is likely to assign to a person in an ad.

In practice, the variable has a string format but a repetitive structure, and it reports the description of each scene sequentially, separated by a “/” symbol. For instance, the Ad Intel database documentation reports “man/woman/auto lot/close up of car” as an example of a creative description. As the example suggests, the creative descriptions are explicit in reporting the gender of the people who appear in the ads, and they map the contents of the ads synthetically but closely. Figure 1 shows an example of the mapping between the creative ad description and the scenes of a beer brand ad.

The Ad Intel data can also be matched with estimated viewership data, which measures the

**Figure 1:** Example of Advertising Creative Description and the Corresponding Scenes in the TV Ad.



number of impressions for each ad. The ad impressions can be characterized by gender at the household level. We capture the estimated viewership for each ad using the impression files from Nielsen Ad Intel. These files provide the total number of households and/or persons who were exposed to the respective ad on a given media channel, within a particular designated market area (DMA), at a given point in time, segmented by demographic sub-groups (e.g., Male 21-24 years old, Female 18-20 years old). For each occurrence of an ad, we estimate the total audience by gender (and the corresponding male or female audience shares).

The Ad Intel dataset and the ad creative description variable provide the basis to quantify the gender role model indicator and the metrics of gender in context.

## 2.2 Homescan Household Panel Data

The Homescan household panel records household-level transactions, purchase quantities, and prices paid for each trip at the product level. We use these data to estimate the extent to which advertised products are bought primarily by men, women, or both. The panel surveys about 60,000 households every year and collects information on the size and composition of the household, income, type of residence, race, and education, along with the age, gender, and employment status of every member of the household.

Throughout the analyses, we will mainly focus on a sub-panel of about 15,000 households composed of single men or single women, because only in these cases can we precisely pinpoint the gender of the principal shopper and the main consumer of the products.

## 2.3 Survey

In addition to the Nielsen datasets, we collected survey data from 693 students at a large university in Asia. The survey ran between March and April 2024 and aims to provide a complementary “gender in context” metric that does not hinge on the presence of a gender role model in the ad. Appendix A reports the detailed outline of the survey and some descriptive statistics, while Section 2.4 explains the survey-based scoring procedure in detail.

## 2.4 Quantifying Gender Representation

We measure gender representation in TV ads by considering both the extent to which men and women appear in the TV ad (gender represented as a “role model”), and the context in which the role model is represented (“gender in context”, “context associations”, or “context scores”).

### 2.4.1 Representation of Gender as Role Models

Numerous studies of gender representation in mass media and advertising have focused on the representation of genders as role models – that is, whether people of certain genders appear in the medium at all (e.g. Aneja et al. 2023, 2024, Ash et al. 2021, Babar et al. 2023, Goli and Mummalaneni 2023, Hartmann et al. 2023, Kim et al. 2023, Kirgios et al. 2022, Overgoor et al. 2023).

We take a similar approach and quantify gender role model representation in TV ads using a keyword-matching procedure. The keyword list we used for this task is reported in Appendix B. The resulting gender role model indicators for ad  $a$  are then calculated as follows:

$$\text{Male}(a) = \begin{cases} 1, & \text{if any male keyword appears in ad } a \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

$$\text{Female}(a) = \begin{cases} 1, & \text{if any female keyword appears in ad } a \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

### 2.4.2 Representation of Gender in Context

Representation in mass media refers both to the numerical analysis of men’s and women’s presence in the medium, and to the conceptual ideas, associations, and impressions about men and women in society (Krijnen and Van Bauwel 2021). Most prior research has systematically characterized genders for their primary role and presence in ads, focusing on the numerical analysis of gender presence in mass media. Only in a few cases has prior research considered the

rest of the visuals, contexts, and words that appeared together with the represented gender. For instance, Ganahl et al. (2003) conduct a content analysis of the role of men and women not only as primary but also as secondary characters in TV ads from three U.S. networks in 1998. In the analysis, Ganahl et al. (2003) find that women were often cast as secondary, younger, supportive counterparts to men in the ads. In a meta-analysis of gender roles in advertising, Eisend (2010) highlights the importance of measuring the “gender in context”. The meta-analysis suggests that several variables related to gender roles in ads, including behavioral and occupational cues, can make content significantly more associated with male or female role models, and sometimes even represent genders in stereotypical ways. These cues include, for instance, representing a man as a truck driver versus a woman as a housewife, or representing men in positions of leadership versus women in positions of childcare. The representation of gender role models in certain contexts can promote gender stereotyping and can lead to “negative consequences that restrict life opportunities, particularly for women” (Eisend 2010).

Therefore, in addition to a well-established characterization of gender representation as genders’ presence in the ads, we propose to characterize the context in which they are represented. To do so, we take two approaches: a data-driven approach, and a survey-based approach.

**Data-driven Context Score.** To calculate a context score from the creative descriptions, we decompose creative description  $a$  into each unique word  $w$  that makes up the description. Then, we calculate a word-level context score,  $R(w)$  as follows:

$$R(w) = \frac{1}{c(w)} \sum_{a=1}^A \sum_{i=1}^{n_a} \mathbb{I}(w_{a,i} = w) \cdot m(w_{a,i}) \quad (3)$$

Where  $n_a$  is the number of words in ad  $a$ ,  $w_{a,i}$  is the  $i^{th}$  word in ad  $a$ ,  $c(w)$  is the total number of times the word  $w$  appears across all ads,  $\mathbb{I}(w_{a,i} = w)$  is an indicator function that is 1 if  $w_{a,i} = w$  and 0 otherwise, and  $m(w_{a,i})$  is a variable with value 1 if the word  $w_{a,i}$  appears in an ad with gendered (e.g. male) role model, and 0 otherwise.

This procedure results in a word-specific “male” or “female” score  $R(w)$  ranging between 0 (the word never appears with a male/female role model) or 1 (the word appears exclusively associated with a male/female role model). To avoid cases with very extreme scores but little support, we only count words that appear at least 3 times over the 2,116,760 ad creative descriptions.

Finally, the ad-level measure  $M(a)$  captures the average male or female context score for ad  $a$  by averaging the  $R(w)$  scores of all words  $w_{a,i}$  used in the ad:

$$M(a) = \frac{1}{n_a} \sum_{i=1}^{n_a} R(w_{a,i}) \quad (4)$$

The difference between the male and female representation scores gives a number between -1 (i.e., the ad includes exclusively words associated with a female role model) and 1 (the ad includes exclusively words associated with a male role model).

The data-driven context score  $M(a)$  in Eq. 4 is statistically straightforward, transparent, and flexible to alternative specifications. However, it also requires the presence of gender role models across the ads to obtain variation in the score. Therefore, we complement this metric with a survey-based score that does not rely on the representation of genders as role models for variation.

**Survey-based Context Score.** We present each survey respondent with 34 ad creative descriptions to label, of which 3 are practice cases and 1 is an example case. We previously removed all the clearly gendered keywords from the ad descriptions and substituted them with a gender-neutral alternative (for instance, all mentions of “man” or “woman” were replaced by “person”, “men/women” by “people”. See Appendix B for the list of keywords). This ensures that the respondents only consider the “context words” and not the gender role models for their labeling task.

Then, we ask respondents whether they think the ad content is most likely associated with a man (or group of men), a woman (or group of women), or if both are equally likely to be associated with the ad content. We base this question on the measurement of stereotypical gender beliefs in Bordalo et al. (2019) and ask respondents to express their opinion on a scale ranging from -1 (most likely associated with man/men) to 1 (most likely associated with woman/women). The mid-range point at 0 indicates that both groups are equally likely to be associated with the content of the ad. Respondents also have the option to label the association as “Not Applicable” (NA) (Bordalo et al. 2019).

This procedure results in a survey-based context score  $H(a) \in \{-1, 1\}, \text{NA}\}$  for each ad  $a$ . We impute the context association score for ads that were scored as “NA” as the average score for the ads in the same industry. The resulting score with imputations is then equal to  $\widehat{H}(a) \in [-1, 1]$ .

The survey respondents scored 18,372 unique ad creative descriptions. We predict the scores of the remaining, unlabelled descriptions using a Ridge regression.<sup>4</sup> To train the Ridge regression and predict unlabeled cases, we stem the ad creative descriptions and convert them

---

<sup>4</sup> $MSE = 0.127, R^2 = 0.482$ , Pearson correlation between test cases and predicted cases = 0.695, p-value < 0.001.

to matrices of TF-IDF features, measuring the relative frequency of stemmed words within an ad, and adjusting for how common or rare the stemmed words are across all ads (Pedregosa et al. 2011).

## 2.5 Validation

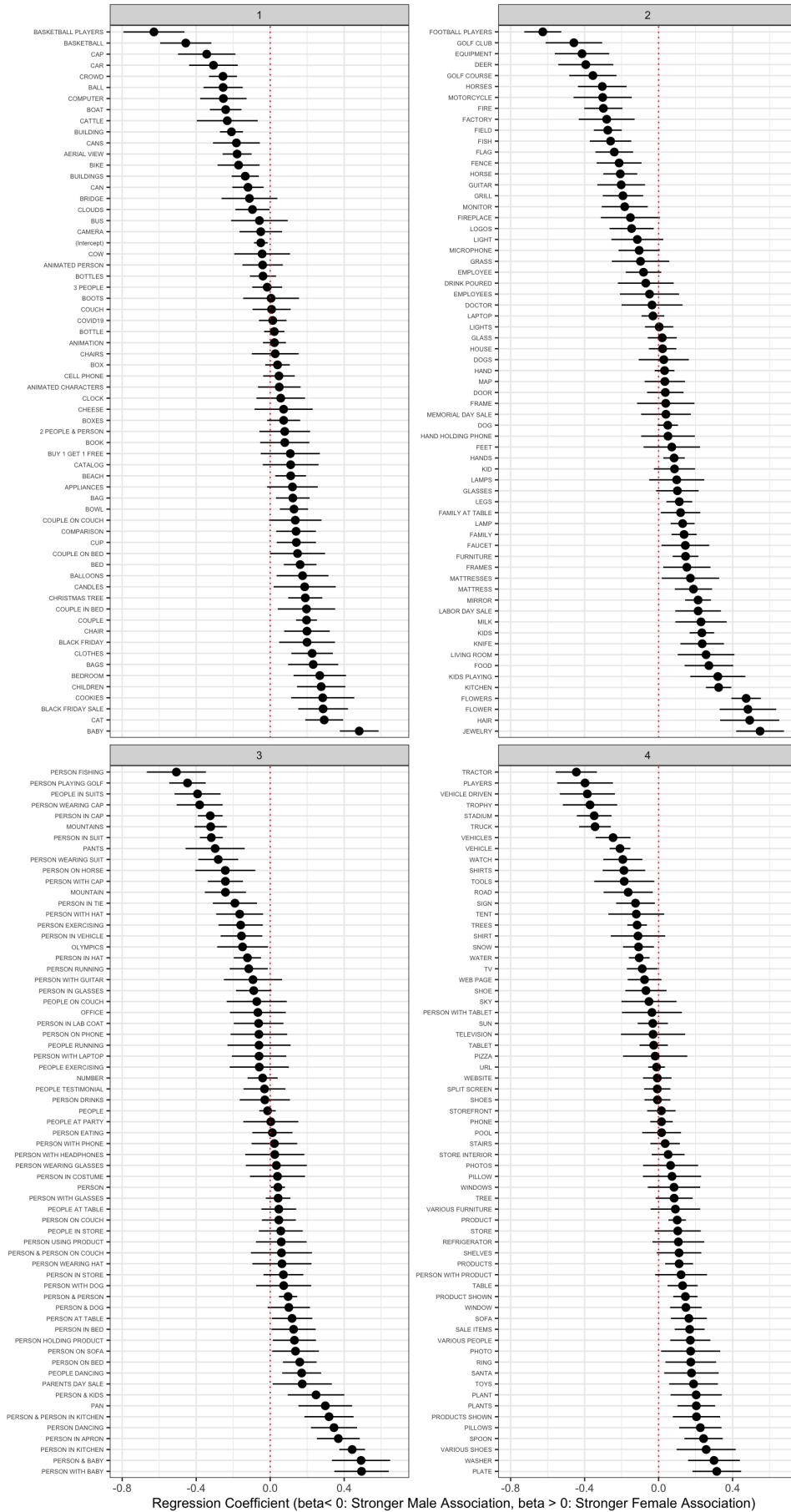
The three metrics for gender role models, data-driven context, and survey-based context have distinct but complementary interpretations. The gender role model metrics are intuitive indicators that advertisers place a person of a given gender in at least one scene of the ad creative (and that a viewer would perceive that presence). The data-driven context score measures the extent to which advertisers tend to place certain words, concepts, and visuals in ads with a person of a given gender. The survey-based context score measures the extent to which viewers would associate words, concepts, and visuals (that are not per se gendered) with a given gender.

The results of the context-scoring procedures suggest that we can capture meaningful associations between ad contexts and gender representation using the context scores. For instance, Figure 2 shows that when we regress the survey-based association score on single ad creative words, the resulting coefficients describe fairly typical gender norms represented in TV ads – such as football and basketball players associated with male representation, and childcare and housekeeping associated with female representation.

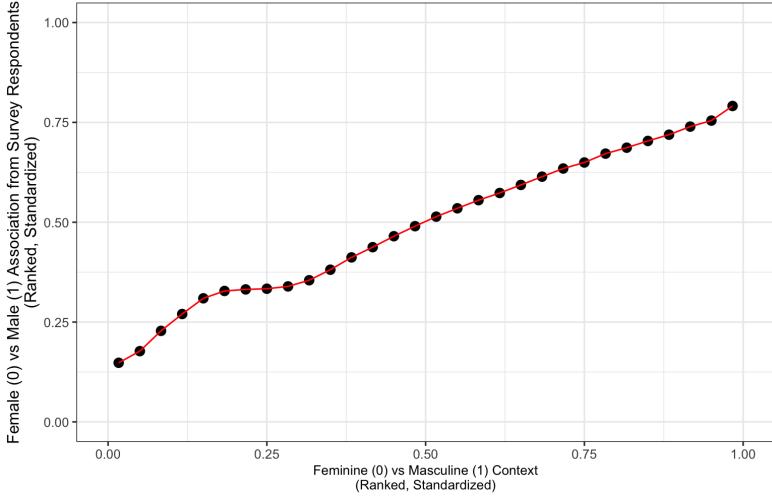
The survey-based association scores are also correlated with the data-driven context scores based on gender role model representation (Pearson's correlation coefficient for ranked and scaled values = 0.642,  $t = 1131.1$ ,  $df = 1827960$ ,  $p\text{-value} < 0.001$ , Figure 3). The correlation between the residual variation in the two metrics is still moderate to strong even after controlling for brand and product fixed effects (Pearson's correlation coefficient for ranked and scaled values = 0.52,  $t = 824.34$ ,  $df = 1827960$ ,  $p\text{-value} < 0.001$ , Appendix Figure 15). This suggests that advertisers may be placing gender role models in contexts that correlate with viewers' associations with the genders represented.

The respondents were also asked to express their beliefs about each male and female knowledge in several categories – including the Kardashians, Disney movies, emotion recognition, mathematics, cars, and sports – to estimate survey participants' baseline beliefs about gender roles (Bordalo et al. 2019). The average scores for these questions reveal that the beliefs of survey participants reflect widespread beliefs and stereotypes about gender roles in society. Furthermore, respondents of different genders scored some knowledge categories in a gender-congruent way – relatively more positively for female respondents, and relatively more negatively for male respondents (Appendix Figure 14).

**Figure 2:** Regression Coefficients: Single Words. Formula:  $H(a) = \alpha + \beta w + \varepsilon$



**Figure 3:** Average Human Association Scores versus Average Context Scores, Ranked and Scaled. Pearson’s correlation coefficient = 0.642, t = 1131.1, df = 1827960, p-value < 0.001.



### 3 Gender Representation in TV ads between 2010 and 2020

In this section, we provide an overview of the extent of gender representation in TV ads between 2010 and 2020. First, we present the average representation of genders as role models in the ads by year and product industry (Section 3.1). Then, we describe the context in which the role models have been represented in the ads (Section 3.2). Next, we assess to what extent any heterogeneity (and imbalance) in representation reflects gender-based targeting of TV audiences or targeting of consumers with desirable marketing potential (Section 3.3). Finally, we highlight some notable cases of imbalance in gender representation that cannot be explained only by TV audience targeting or by patterns of historical consumption (Section 3.4).

#### 3.1 Female and Male Role Model Representation in TV Ads

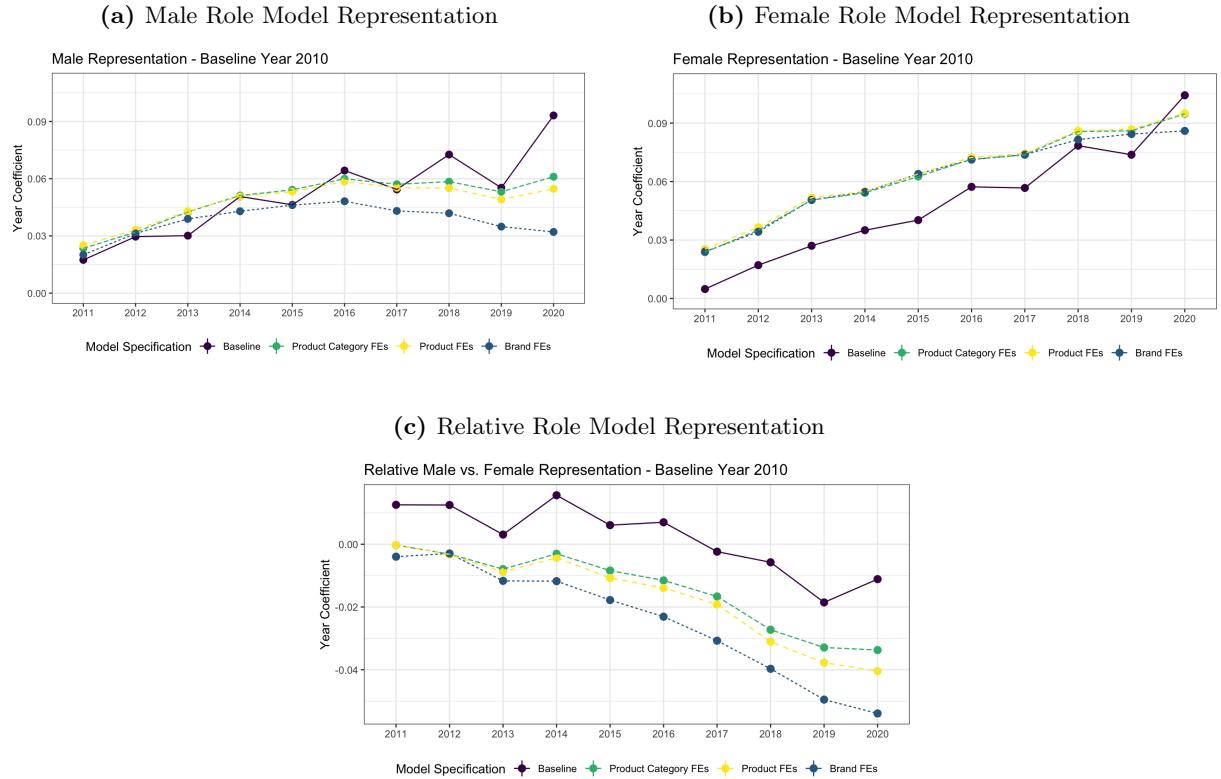
Recent research on gender representation in TV programs highlighted how women are generally less represented than men in cable news and how they are also assigned to discuss different topics than men (Goli and Mummalaneni 2023). Observing the representation of gender role models over time allows us to assess whether this pattern also interests men’s and women’s presence in TV ads.

Figure 4 describes the extent of representation of male and female role models in TV ads every year relative to 2010. The years indicate the first time ads were broadcasted. Figure 4 shows a baseline upward trend towards more representation of both men and women in TV ads. The percentage of ads representing women and men as role models increased respectively from 27.9% and 36.2% (2010) to 38.3% and 45.5% (2020). Therefore, advertisers have chosen to prominently represent both men and women in TV ads more frequently over time – instead

of, perhaps, showing products, objects, backgrounds, or text.

Average female role model representation, starting at 27.9% in 2010, has steadily increased consistently even within brands, product categories, and products – with a more pronounced increase between 2017 and 2020 (Figure 4b). On the other hand, male role model representation increased between 2011 and 2016, then plateaued or decreased between 2016 and 2019 relative to the past, especially when considering ads within brands and product categories. Average male role model representation in the ads has increased again in 2020 compared to the baseline (Figure 4a). Overall, Figure 4c indicates a downward trend in the representation of men relative to women as role models in the ads. In particular, 2017 was the first year in which the relative representation of women as role models in TV ads exceeded that of men compared to the baseline.

**Figure 4:** Yearly Representation of Gender Role Models across TV Ads (2011-2020, baseline year 2010). Colored lines indicate linear models with fixed effects for brands (blue), product categories (green), and products (yellow). Regression specifications for (4a) and (4b) are of the form  $\text{Male}(a) = \alpha_{\text{year}} + \varepsilon | \text{Fixed Effects}$  and  $\text{Female}(a) = \alpha_{\text{year}} + \varepsilon | \text{Fixed Effects}$ . Regression specification for (4c) is of the form  $\text{Male-Female}(a) = \alpha_{\text{year}} + \varepsilon | \text{Fixed Effects}$ .

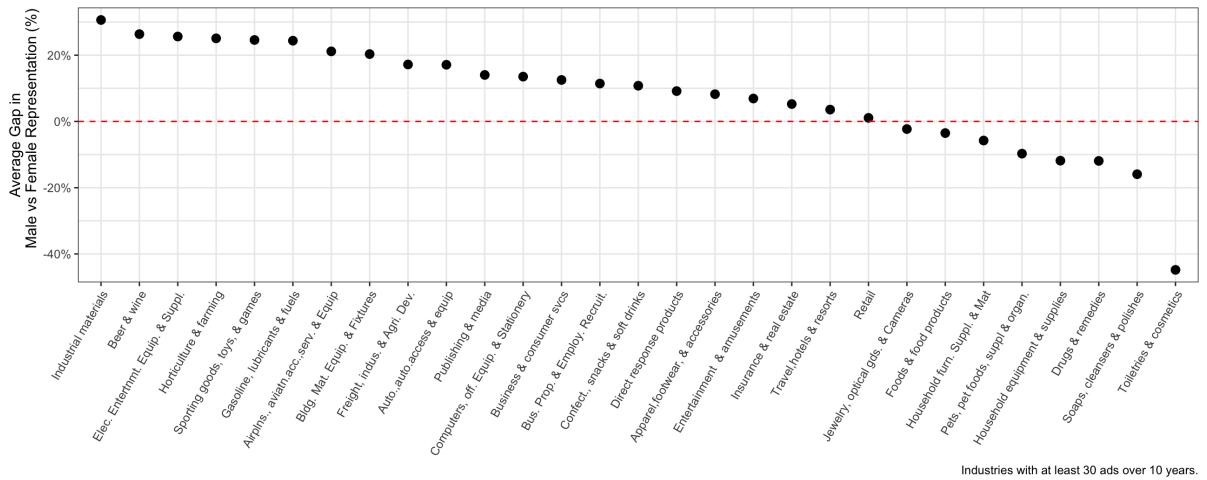


These patterns in role model representation could be heterogeneous across industries. Considering industries with at least 30 ads, Figure 5 shows that women have appeared as role models relatively more than men in only 8 of 29 industries: toiletries and cosmetics; soaps, cleansers, and polishes; drugs and remedies; household equipment and supplies; household fur-

niture, supplies, and materials; pets, pet foods, supplies, and organization; and jewelry, optical goods, and cameras. Most of these industries are related to traditional gender roles based on the International Social Survey Programme (e.g., physical appearance, roles of care, and household upkeep; Bordalo et al. 2016, ISSP 2016).

On the other hand, men have appeared as role models relatively more than women in 21 industries and approximately the same in 1 industry (retail). Some of these industries are also related to traditional gender role models for men. For instance, the over-representation of male role models in business & consumer services ads may be related to the traditional belief that “a man’s job is to earn money, a woman’s job is to look after the home and family” (ISSP 2016). Other industries in which men are over-represented compared to women include sporting goods; beer & wine; automobiles & auto accessories; apparel, footwear & accessories; and computers, office equipment & stationery. This evidence is suggestive of the fact that TV ads may contribute to perpetuating traditional beliefs about gender role models.

**Figure 5:** Average Relative Representation of Male (versus Female) Role Models in TV Ads per Product Industry (2010-2020, industries with at least 30 ads). The dashed red line indicates equal representation.



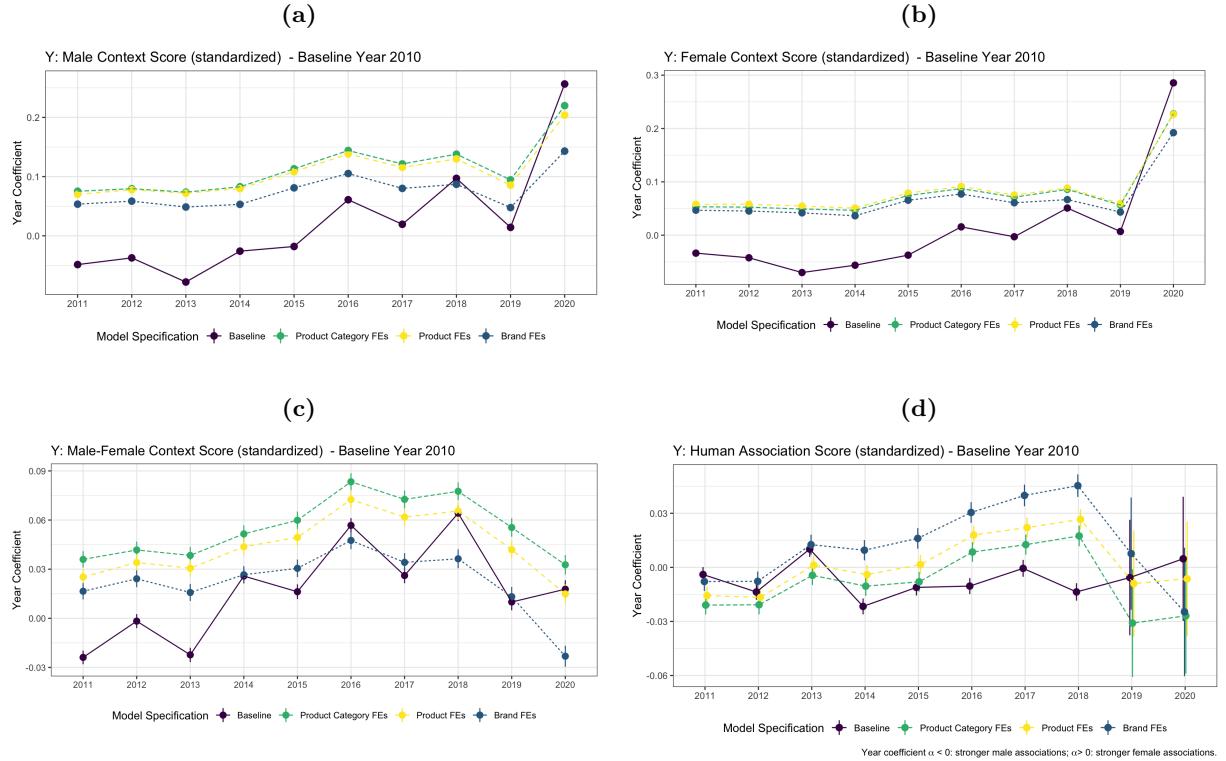
### 3.2 In Which Contexts Have Men and Women Been Represented in TV Ads?

The heterogeneity in gender representation across industries suggests that the context in which men and women are represented in TV ads is important for understanding the relationship between genders as role models and the representation of their roles in society. The representation of genders in context might reinforce traditional gender roles and potentially perpetuate gender stereotypes, above and beyond role model representations.

Figure 6 demonstrates the yearly average context scores, which characterize the extent to which words and concepts in the TV ads were mostly associated with men and women. Figures

6a – 6c, which refer to the data-driven context scores, show that TV ads included more words and concepts associated with the presence of both men and women over the years, with a sharper increase of female vs. male contexts between 2019 and 2020 compared to the baseline year. Figure 6d, which refers to the perceptions of survey respondents independently of the presence of gender role models, shows that, within brands and products, TV ads have included relatively more words and concepts associated with women over time.

**Figure 6:** Yearly Representation of Context Associations across TV Ads (2010-2020 – baseline year 2010). Colored lines indicate linear models with fixed effects for brands (blue), product categories (green), and products (yellow). Regression specifications for 6a and 6b are of the form  $M(a) = \alpha_{\text{year}} + \varepsilon | \text{Fixed Effects}$  where  $\tilde{M}(a)$  is the standardized and centered data-driven context score for male and female contexts respectively. Regression specification for 6c is of the form  $\tilde{M}(a) = \alpha_{\text{year}} + \varepsilon | \text{Fixed Effects}$  where  $\tilde{G}(a) = M(a)_{\text{Male}} - M(a)_{\text{Female}}$  is the standardized and centered data-driven context score for male relative to female contexts. Regression specification for (6d) is of the form  $\tilde{H}(a) = \alpha_{\text{year}} + \varepsilon | \text{Fixed Effects}$ , where  $\tilde{H}(a)$  is the standardized and centered survey-based context association score.



Next, we assess the relationship between the presence of gender role models in ads and the contexts in which they are represented. Male and female role models have generally been represented in gender-congruent contexts over the last decade (Appendix Figure 16). In other words, TV ad creatives in which women (men) were represented as role models also included words or concepts most frequently associated with women (men). The results also suggest interesting patterns for gender representation in context between 2018 and 2020. From 2018 onward, within brands, products, and product categories, both men and women were represented

as role models in relatively less gender-congruent contexts. That is, TV ad creatives have recently placed women (men) in contexts that viewers might associate with the opposite gender.

Figure 7 describes the average gender representation in context across product industries using the survey-based context scores (see Appendix Figure 17 for the data-driven context scores). On average, in 22 product industries, when no role model is present, ads represent contexts generally more associated with men. Furthermore, the presence of genders as role models in TV ads is associated with gender-congruent representations in many industries. For example, in toiletries & cosmetics, retail, household upkeep, and pet supplies ads, when a woman is represented as a role model, the rest of the context is also much more likely to exhibit words and concepts associated with women. On the other hand in sporting goods, toys & games ads, and beer & wine ads, when a man is represented as a role model, the rest of the context includes more words and concepts associated with men. As discussed earlier, also in these cases, advertisers seem to represent the role models of different genders in contexts congruent with traditional gender roles (ISSP 2016).

Notably, in several cases, the presence of a role model either has no particular contextual associations to genders or even enhances context associations with the opposite gender. For instance, across soaps, cleansers & polishes, and household upkeep ads, male role models are represented in more female contexts. Instead, across automotive, auto accessories & equipment products, gasoline, lubricants & fuel products, industrial materials, and computer products, women are represented in ads in more male contexts, even when they appear as role models.

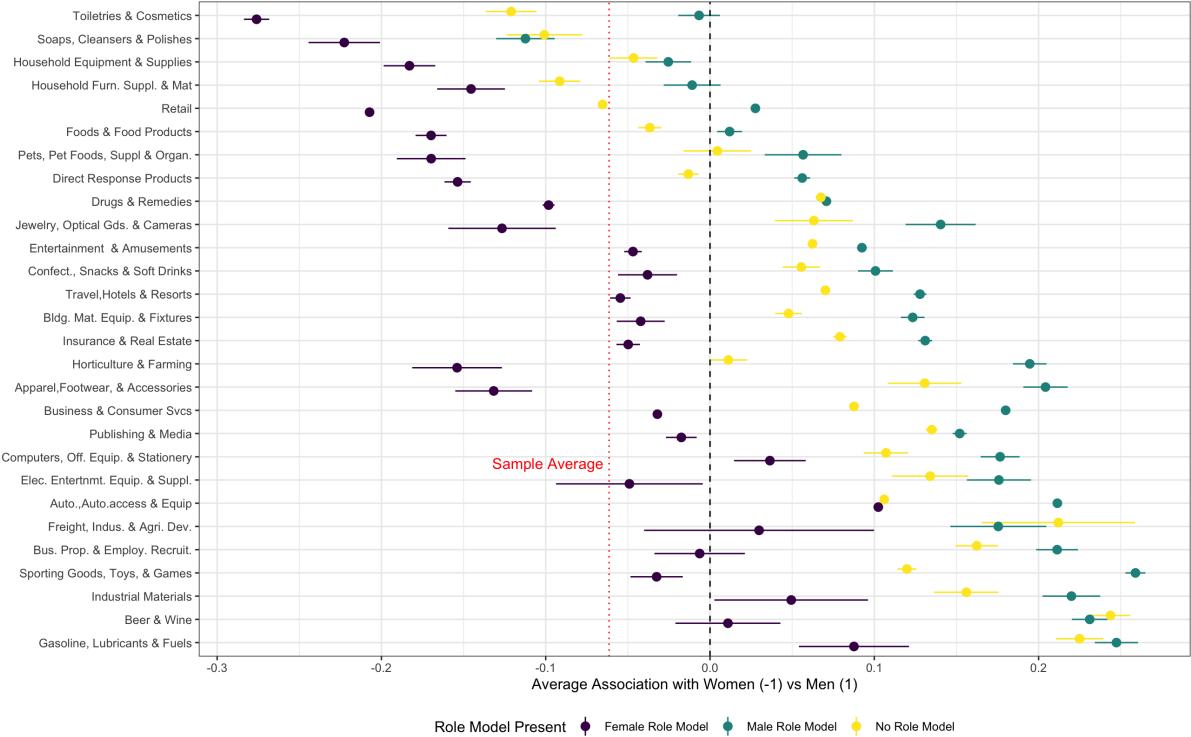
In Section 3.3, we explore the extent to which gender representation in TV ads may be related to the targeting of TV audiences and/or consumers' spending patterns in the product categories.

### **3.3 Using Gender Representation to Connect with Target Audiences and Consumers**

A question that may arise from the analyses above is whether advertisers in different industries are resorting to gender role models and genders in context to connect with the consumers who are most likely to watch the TV program in which the ads were aired.

To address this question, for each product with a minimum of 30 ads, we measure the correlation between the weighted average male vs. female survey-based context score and the weighted average male audience share for the corresponding ads ( $N=976$ ). The averages are weighted by the ad's total audience to capture systematic differences in advertising reach. Figure 8 shows a positive correlation between male vs. female survey-based context score and male

**Figure 7:** Yearly Representation of Role Models and (Survey-based) Context Association across TV Ads and Product Industries. Colors indicate the presence of a male (yellow), female (green), and no role model (purple) represented in the ad. The dark red points indicate the sample average survey-based association scores per industry. The figure includes industries with at least 100 ads over 10 years.



audience share ( $r=0.639$ )<sup>5</sup>. This suggests that advertisers may use gender representation to connect with an intended target audience – although the correlation is weaker at higher levels of male viewership.<sup>6</sup>

Given a target audience, however, there is still unexplained residual variation in male vs. female context scores. For instance, while ads about luggage, shotguns, dance classes, and rifles are targeted at a similar share of male viewers, luggage and dance classes ads are systematically displayed in more female contexts, while shotguns and rifles ads are systematically displayed in more male contexts.<sup>7</sup>

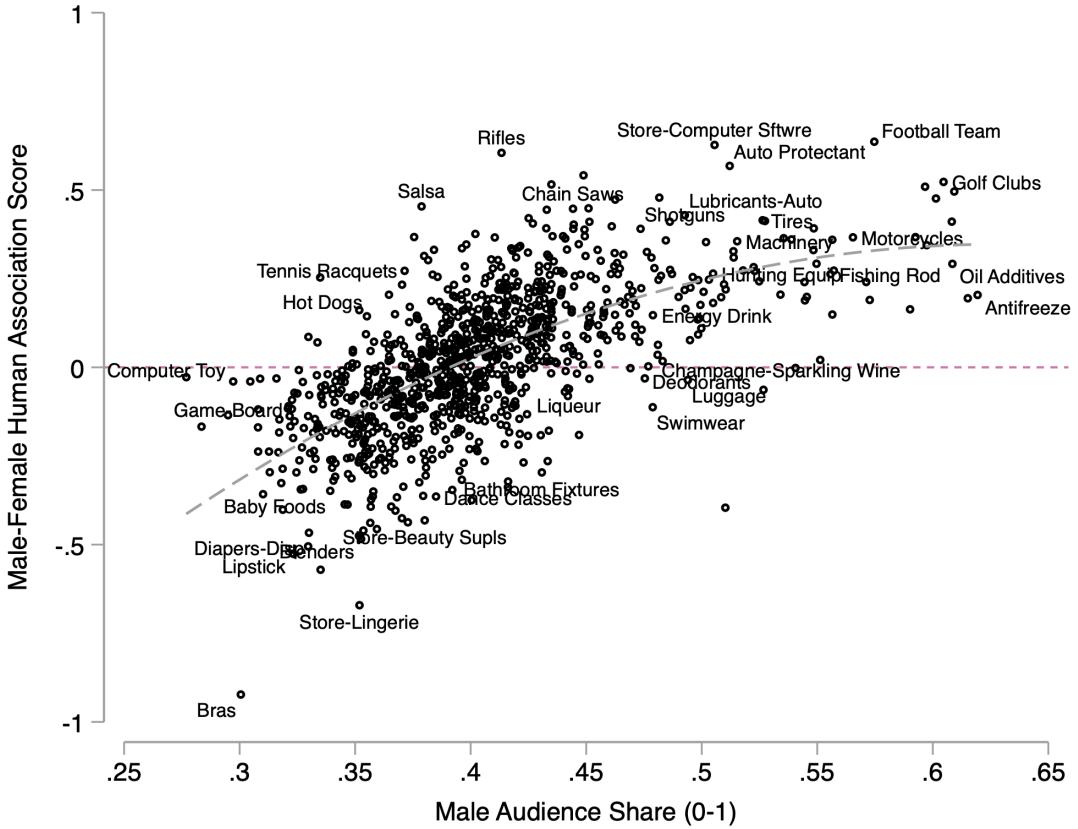
The unexplained residual variation in gender representation may reflect differences in product purchase decisions by consumers of different genders, beyond target TV audiences. For instance, bras or lipsticks are almost exclusively purchased by women, while golf clubs remain mostly purchased by men. Therefore, advertisers might reasonably resort to gender-congruent

<sup>5</sup>A similar pattern arises if we look at role model representation instead of gender in context ( $r=0.578$ ,  $N=976$ ). The correlation at the industry level is also higher ( $r=0.857$ ,  $N=32$ ), as evidenced in Appendix Figure 19.

<sup>6</sup>If audience targeting is primarily done at the brand level, one would expect a higher brand-level correlation. However, using the same procedure but aggregating by brands instead of products, we find a much weaker correlation ( $r=0.235$ ,  $N=12,736$ ).

<sup>7</sup>Similar results are found in terms of role model representation. For instance, hunting equipment ads have a 71% male representation (vs. 14% female representation), while their viewership is gender-balanced.

**Figure 8:** Average Male vs. Female Survey-based Association Scores versus Average Male Audience Share by Products. Averages are weighted by the ads' total audience. Pearson's correlation coefficient = 0.639 (p-value<0.001, N=976)



representations to connect with their customers in these product categories, above and beyond the demographics of their TV audiences.

To assess this empirical question, we exploit the yearly Homescan household panel of transactions, purchase quantities, and prices paid for each trip at the product level, focusing on single men and single women households. The AdIntel dataset and the Homescan household panel do not use the same product identifiers. Therefore, we identify 60 widely consumed products that can be matched across both datasets. Together, these products account for 60% of the total yearly expenses of single households. With these data, we run a series of Ordinary Least Squares (OLS) regressions of the following form:

$$\ln(Exp_{it}^p) = \alpha + \beta^p SingleMen_i + Controls_{it} + State_i + Time_t + \epsilon_{it} \quad (5)$$

where  $\ln(Exp_{it}^p)$  is the log of personal expenditure on product  $p$  by consumer  $i$  at time  $t$ ,  $SingleMen_i$  is a dummy equal to 1 if consumer  $i$  is a single man (versus a single woman),  $Controls_{it}$  is a vector of consumer controls for income, age, race, type of residence, occupation,

and education,  $State_i$  is a vector of state fixed effects, and  $Time_t$  is a vector of year fixed effects. The “gender consumption gap” coefficient  $\beta^p$  captures the log difference in expenditure between single man and single woman for product  $p$ , holding everything else constant. In other words,  $\beta^p = -0.25$  means that (single) men spend on average 25% less on product  $p$  than (single) women, holding everything else constant. In the models, we apply the population weights provided by Nielsen to make purchase figures representative at the national level.

Figure 9 plots the correlation between the (weighted) average product-level male vs. female survey-based context scores and the corresponding  $\beta^p$  coefficient estimated using Equation (5). Similar to the case of audience shares, we find a positive correlation between the male vs. female survey-based context scores and the gender consumption gaps ( $r=0.612$ ).<sup>8</sup> This correlation is consistent with advertisers of various products attempting to connect with their consumer base through congruent gender representations. For instance, cosmetics, skincare preparations, or hair care products are mostly purchased by women, while automotive-related products, beer, or shaving products are mostly purchased by men. In these product categories, we observe that advertisers rely on gender-congruent representations. However, we also notice important variations. For instance, men spend about 25% less on horticulture and farming products compared to women. However, horticulture and farming ads systematically represent disproportionately male contexts.

### 3.4 Beyond Connection: Gender Misrepresentation

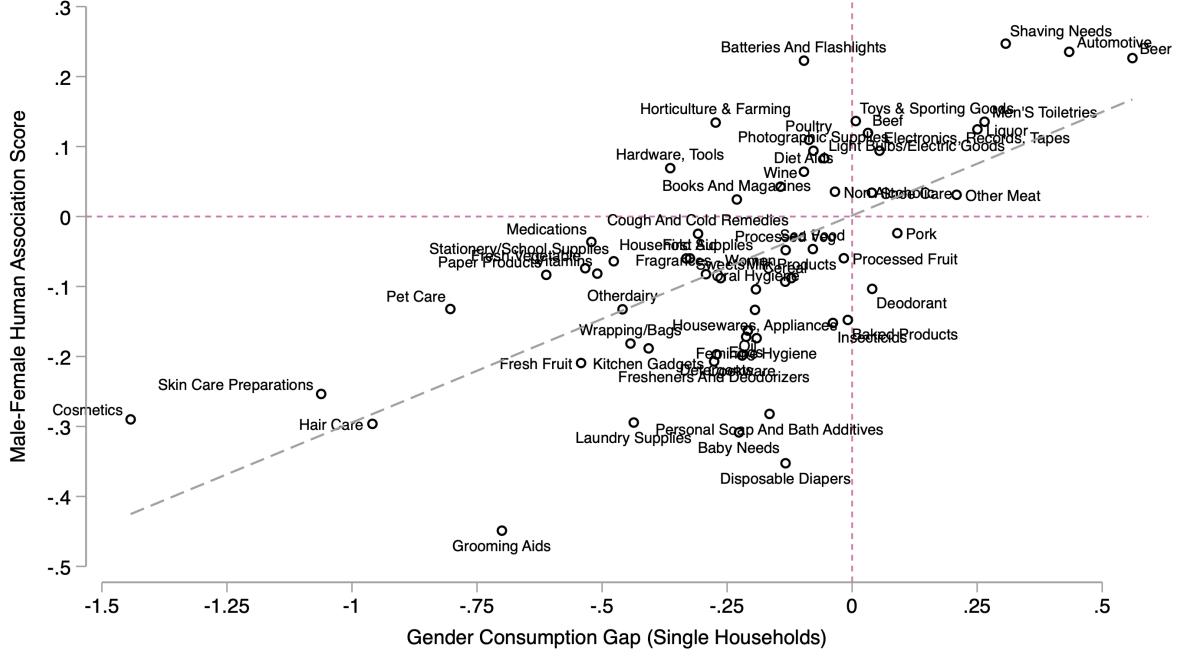
How much residual unexplained variation in gender representation remains after controlling for i) consumption and/or ii) target audience? We standardize each measure of gender representation (i.e., role models and contexts) using z-scores, and regress them on the gender gap in consumption and male audience share respectively.

Figure 10 plots the mean residuals in gender representation for each of the 60 products relative to what the consumer base would predict. The graph reveals that gender representation used by advertisers to promote products such as vitamins, cookware, shaving needs, or detergents is generally aligned with the demographics of the consumer base. However, ads for most personal hygiene products (soap, deodorant, grooming aids), baby products (diapers, baby needs), as well as housewares and laundry supplies, include more female representations than what the consumer base demographics would suggest. In contrast, ads for farming products, books, red meat (beef, other meat), or tools exhibit more male representations – more than half a standard deviation away from what pure connection with their consumer base would predict.

---

<sup>8</sup>Using the average gaps in male vs. female role model representation instead of the survey-based context scores results in a correlation of 0.827.

**Figure 9:** Average Male-Female Survey-Based Association Scores versus Gender Consumption Gap by Products. Average context scores are weighted by ads' total audience, and consumption is weighted by population weights. The consumption gap corresponds to the  $\beta^p$  coefficient estimated from the OLS regression (5). Pearson's correlation coefficient = 0.646 (N=60). These products account for 60% of consumers' reported yearly expenditures.



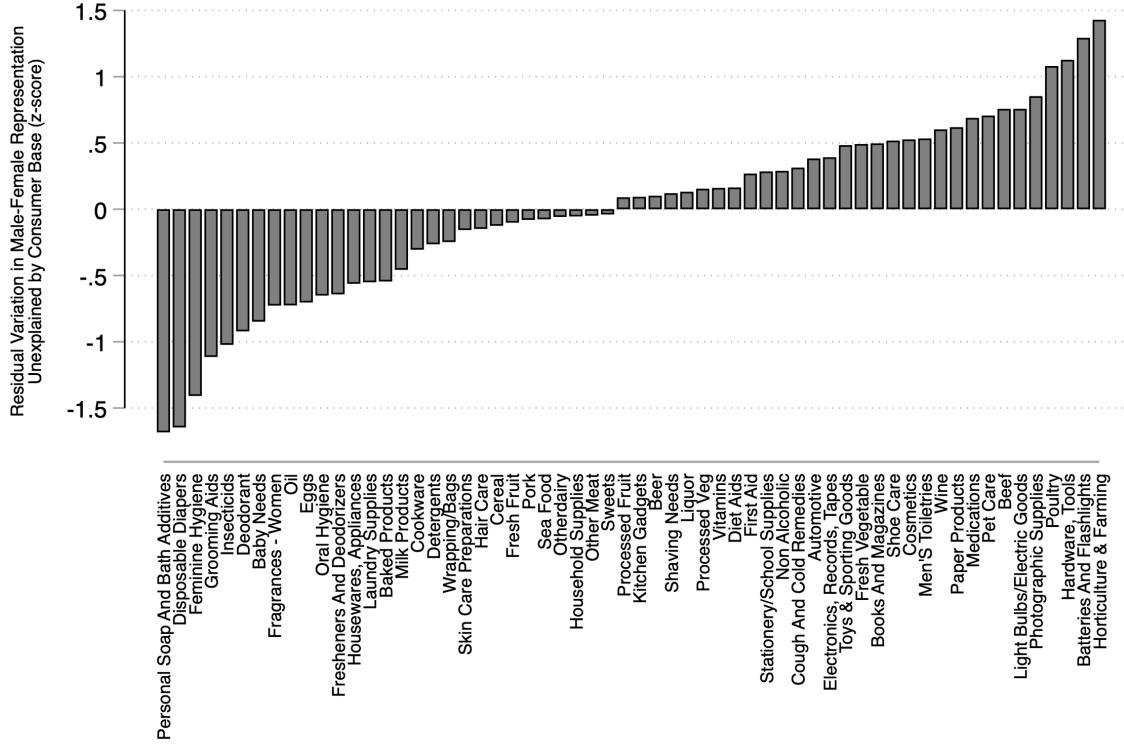
The deviation of representation from the consumer base is highly correlated to deviation from viewership ( $r=0.498$ ), as evidenced in Appendix Figure 18. However, we see some notable differences. For instance, skincare ads are watched by more women than their gender representation would predict, but the representation is aligned with the consumer base. Instead, photographic supplies are represented in a more male context than what their consumer base would predict, but the representation is aligned with their male audience share.

## 4 Discussion

Our analyses provide a complex and nuanced picture of gender representation in TV advertising over the past decade. We document a general alignment between how brands represent gender across industries and products, and the intended target audiences. However, we also reveal major disconnects in several product categories or industries between gender representations and the demographics of intended target audiences and consumer bases. In this final section, we discuss potential explanations for these discrepancies.

Figure 10 illustrates the male-female consumption gap over time alongside gender representation in ads for products with varying degrees of disconnect from the consumer base: no

**Figure 10:** Deviation of gender representation from connection with the consumer base. Each measure of gender representation (as role models, and in context) is standardized using z-scores and regressed on the gender gap in actual purchase behavior. The Figure plots the mean residual variation of both regressions.



deviation (e.g., automotive), male deviation (e.g., horticulture/farming), and female deviation (e.g., deodorant, disposable diapers). The consumption gap is estimated annually using Eq. (5), and gender representation is measured through a composite index (as role models, and in context).<sup>9</sup>

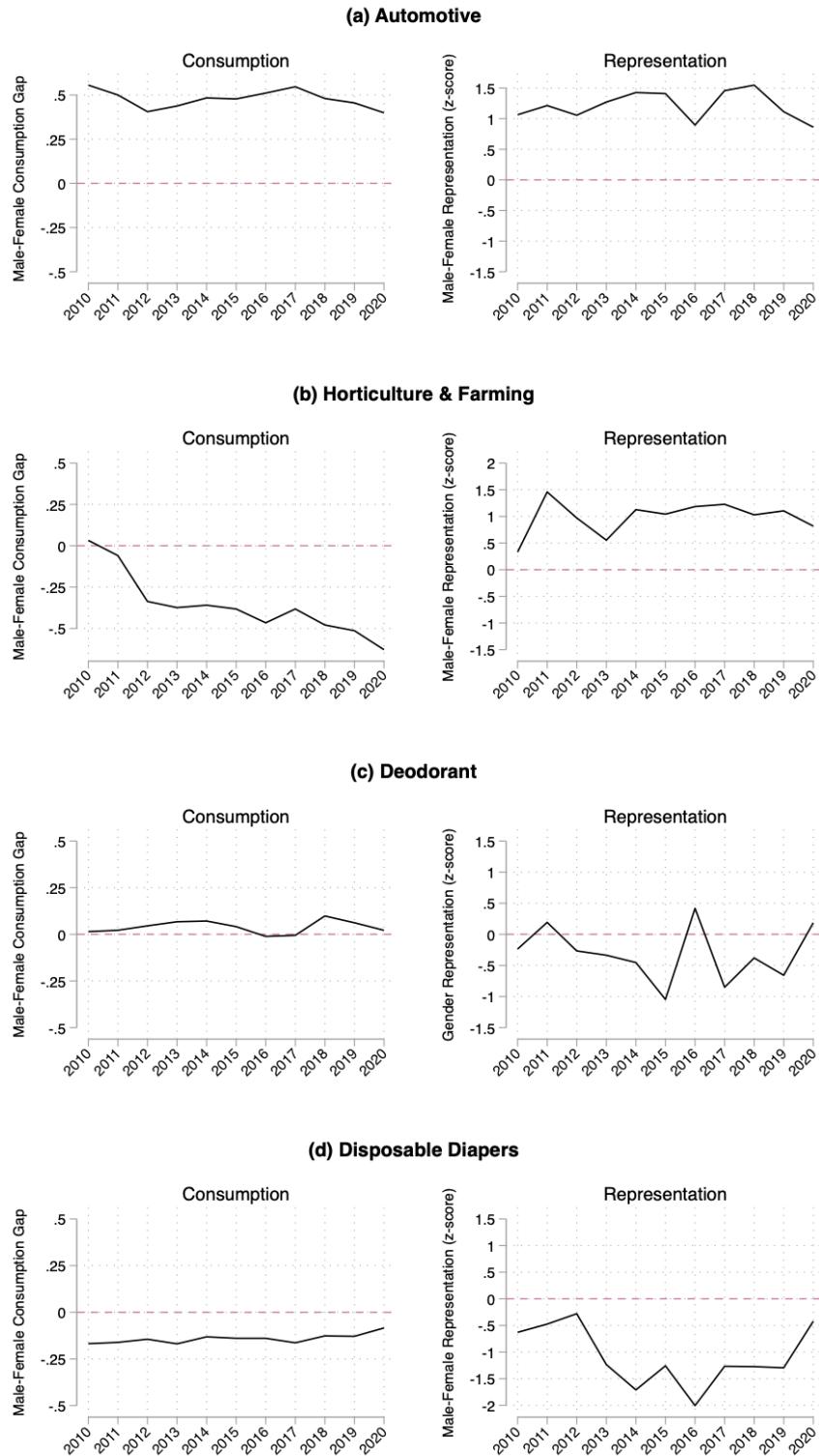
Figure 11a suggests, for instance, that men consistently outspend women in the automotive category, which translates into a stable male representation in advertisements. This indicates gender-based targeting of marketing messages based on historical spending patterns.

In the horticulture and farming products category, however, we observe a shift towards increased female spending over time, yet the gender representation remains predominantly male (Figure 11b). This pattern illustrates a first possible explanation: advertisers in certain sectors may persist in targeting a traditionally male or female audience, potentially underestimating or not fully internalizing demographic and consumption shifts. This is possibly due to defaulting to past habits in the absence of data-driven insights, or more general forms of marketing inertia.

A second explanation for the deviations involves repositioning strategies to attract new

<sup>9</sup>The composite measure is derived by z-scoring each measure and averaging the two z-scored indices.

**Figure 11:** Male-female consumption gap vs. male-female gender representation, selected products. The male-female consumption gap is estimated by panel year using equation (5). Gender representation (as role models, and in context) is standardized using z-scores and the mean z-scored value of both measures is plotted over time.



customer segments. For instance, men and women spent approximately the same amount on deodorants between 2010 and 2020 with no significant trend. However, in several years, brands shifted towards more female representation, possibly to attract female customers (Figure 11c). The use of gender representation for (re)positioning purposes can be further illustrated by looking at brand-level gender representations in product categories that do not necessarily lean towards a male or female audience. For instance, within the liquor category, various brands chose to represent gender in different contexts, from female (e.g. Baileys, Grand Marnier) to neutral (e.g. Tuaca) or male (e.g. Drambuie, Jagermeister; Figure 12a).

Finally, other systematic deviations may be consistent with the propagation of old clichés or gender stereotypes. For instance, disposable diapers show a progressive closure of the gender consumption gap, from 17% in 2010 to 8% in 2020. However, gender representation in this product category has been increasingly associated with women until 2016, with some reversal between 2017 and 2020. In these cases, brands may be reluctant to depart from prevailing gender norms, and only a few brands may try to reposition to more neutral associations. Within the diapers category, all major brands resort to very female representations in their ads (e.g., Pampers, Huggies, Luvs), and the few brands that adopt slightly less female representations are relatively niche brands (Figure 12b).<sup>10</sup> Similarly, all grill brands adopt predominantly male representations in their advertising, while no grill brands choose a neutral or female representation (Figure 12c).

## 5 Conclusions

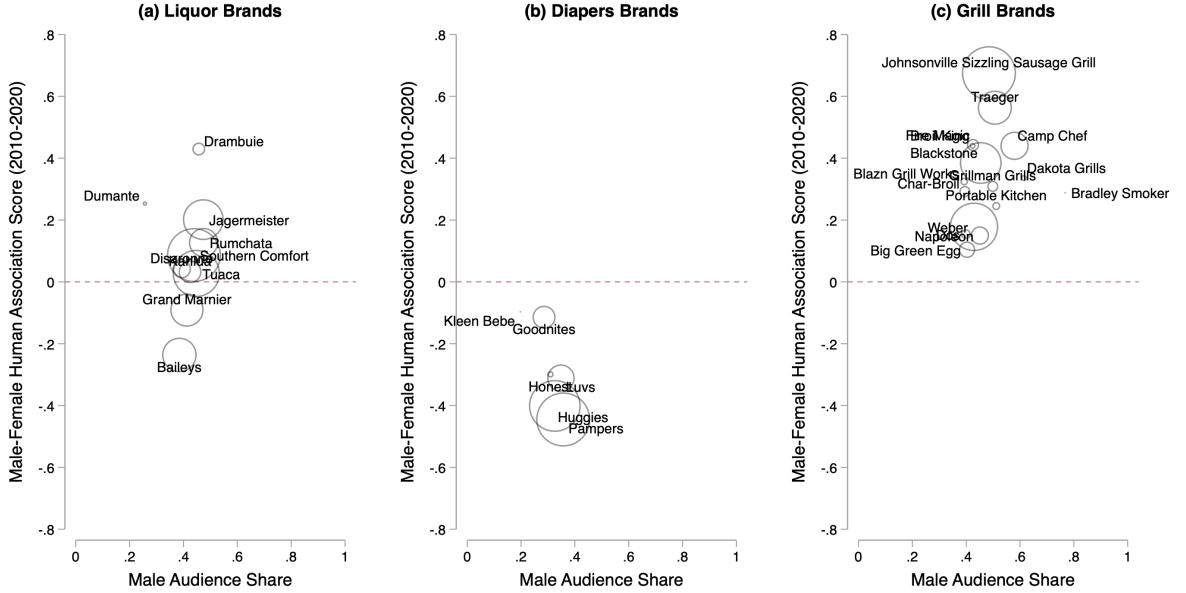
This paper illustrates how advertisers have represented men and women in their TV ads between 2010 and 2020. We propose several metrics of gender representation that take into account both the presence of men and women as role models in the ads and the context in which they are represented. The study reveals a nuanced evolution of the representation of gender roles in TV ads. Women's representation as role models increased overall, but men appeared more frequently as role models across a broader range of industries. Notably, women were predominantly featured in roles and contexts that align with traditional gender roles, particularly in industries like toiletries, cosmetics, and household supplies. Conversely, men were overrepresented in sectors such as sporting goods and automobiles, which are also traditionally male-dominated (ISSP 2016).

The context in which genders are portrayed has also evolved, with TV ads increasingly incorporating female words and concepts, especially within brands. However, there has been a

---

<sup>10</sup>Looking at human associations by year, we see that major Diapers brands do not significantly differ from each other in how they represent genders over time (Appendix Figure 20).

**Figure 12:** Brand-level Male-Female Human Association Score vs. Male Audience Share, Selected Products (2010-2010). Area of circles is proportional to total audience size.



notable recent shift towards more frequent cross-gender contextual representations. It remains unclear whether this change promotes more equal gender representations in TV ads. Across ads for soaps, cleansers, polishes, and household upkeep products, male role models are represented in more female contexts. Conversely, across ads for automotive, gasoline, industrial materials, and computer products, women are represented in more male contexts, even when they appear as role models. While, on paper, this is a signal of cross-gender contextual representation, the associations between roles and contexts still suggest the propagation of traditional gender roles (and potentially harmful stereotypes) in specific industries.

The results also suggest that gender representation in TV ads does not always align with the demographic composition of the target audience or with the consumption patterns of consumers of different genders. Some products consumed overwhelmingly by women or men, such as bras and golf clubs, show expected gender-congruent representations. Other products, like deodorants and horticulture goods, display significant deviations from what gender-based targeting would predict. These inconsistencies suggest that factors beyond consumer demographics, such as attempts to reposition into new customer segments or adherence to stereotypical gender beliefs, might play a role in explaining different gender representations in TV advertising.

This work contributes to the ongoing research on gender representation in mass media and TV advertising. The study also demonstrates the value of using unstructured text from widely employed data sources to study gender representation and context. In the past, researchers have

primarily relied on computer vision for such tasks, with results limited to the quantification of gender role models in mass media.

A limitation of this study is that text captures only a limited number of TV ad features, and that there may be other unobservable (visual) features that the unstructured text cannot capture. We encourage future research to explore ways to combine unstructured text as a data source with visual elements, and combine our metrics with metrics generated by computer vision for a more comprehensive account of gender representation in TV ads. Another limitation is that the survey respondents are not representative of the same population included in the Nielsen datasets, and that the sample size is relatively small. A solution to this problem would be to run additional surveys to measure association scores on larger and more representative samples.

Quantifying gender representation in advertising can foster numerous future research directions and empirical applications. Future research could expand the array of methodologies used to quantify the metrics of gender representation. For instance, an additional context score metric could incorporate insights from generative AI tools and Large Language Models.

Empirical applications focused on temporal or geographical changes in gender representation, particularly considering evolving consumer tolerance of gender stereotypes, are also particularly promising. There is limited knowledge about whether advertisers respond to evolving attitudes and behaviors related to gender representation, and whether these responses influence consumer attitudes or purchase behaviors. For example, while the prevalence of gender stereotypes is well-documented (Coffman et al. 2021, Fischbacher et al. 2023), it has remained relatively stable over the past few decades (Bertrand 2020, Eagly et al. 2020). However, the rise of digitally native social movements opposing systemic discrimination and stereotypes poses several challenges to brands and advertisers.

Recent causal evidence indicates that the #MeToo movement led to a decline in the purchase of stereotypically “female” products, reflecting a shift in consumer tolerance towards gender stereotypes (Bellet et al. 2024). It is unclear whether advertisers should respond to societal shifts and rely on less gender-stereotypical content, or whether the status quo should prevail. A thorough investigation of this phenomenon may build on geographical variation in exposure to gender-stereotypical advertising and its impact on brand-level sales (Kim et al. 2023).

Finally, future work may distinguish cases where backlash is more or less likely to occur, either due to a stronger disconnect between the gender representation and the (local) consumer base of the brand, or to how gender representation is perceived by consumers holding various political views or stereotypical beliefs.

## References

- Aneja, Abhay, Michael Luca, and Oren Reshef**, “The Benefits of Revealing Race: Evidence from Minority-owned Local Businesses,” Technical Report, National Bureau of Economic Research 2023.
- \_\_\_\_\_, **Oren Reshef, and Gauri Subramani**, “Attrition and the Gender Patenting Gap,” *Review of Economics and Statistics*, 2024, pp. 1–31.
- Ash, Elliott, Ruben Durante, Maria Grebenschikova, and Carlo Schwarz**, “Visual representation and stereotypes in news media,” 2021.
- Authority, Advertising Standards**, “Harm and Offence: Gender stereotypes,” 2022.
- Babar, Yash, Ali Mahdavi Adeli, and Gordon Burtch**, “The effects of online social identity signals on retailer demand,” *Management Science*, 2023, 69 (12), 7335–7346.
- Balakrishnan, Maya, Jimin Nam, and Ryan W Buell**, “EXPRESS: Differentiating on Diversity: How Disclosing Workforce Diversity Influences Consumer Choice,” *Production and Operations Management*, 2024, p. 10591478241239934.
- Bellet, Clement, David Dubois, and Frédéric Godart**, “Do Consumers Respond to Social Movements? Evidence from Gender-Stereotypical Purchases After # MeToo,” *Management Science*, Forthcoming, 2024.
- Bertrand, Marianne**, “Gender in the twenty-first century,” in “AEA Papers and Proceedings,” Vol. 110 2020, pp. 1–24.
- Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer**, “Stereotypes,” *The Quarterly Journal of Economics*, 2016, 131 (4), 1753–1794.
- \_\_\_\_\_, \_\_\_\_\_, \_\_\_\_\_, and \_\_\_\_\_, “Beliefs about gender,” *American Economic Review*, 2019, 109 (3), 739–773.
- Coffman, Katherine B, Christine L Exley, and Muriel Niederle**, “The role of beliefs in driving gender discrimination,” *Management Science*, 2021, 67 (6), 3551–3569.
- Daniel, Carlton**, “Harmful Gender Stereotypes or a Double Standard? Calvin Klein and GIRLvsCANCER Ads Banned in the UK,” 2 2024.
- Eagly, Alice H, Christa Nater, David I Miller, Michèle Kaufmann, and Sabine Sczesny**, “Gender stereotypes have changed: A cross-temporal meta-analysis of US public opinion polls from 1946 to 2018..,” *American psychologist*, 2020, 75 (3), 301.
- Eisend, Martin**, “A meta-analysis of gender roles in advertising,” *Journal of the academy of marketing science*, 2010, 38, 418–440.
- Fischbacher, Urs, Dorothea Kübler, and Robert Stüber**, “Betting on Diversity—Occupational Segregation and Gender Stereotypes,” *Management Science*, 2023.
- Ganahl, Dennis J, Thomas J Prinsen, and Sara Baker Netzley**, “A content analysis of prime time commercials: A contextual framework of gender representation,” *Sex roles*, 2003, 49, 545–551.

- Goli, Ali and Simha Mummalaneni**, “Gender Diversity on Cable News: An Analysis of On-Screen Talent and Viewership,” *Available at SSRN 4462592*, 2023.
- Hartmann, Jochen, Oded Netzer, and Rachel Zalta**, “Diversity in advertising in times of racial unrest,” *Available at SSRN 4466378*, 2023.
- ISSP**, “International social survey programme: Family and changing gender roles IV,” 2016.
- Khan, Uzma and Ajay Kalra**, “It’s good to be different: How diversity impacts judgments of moral behavior,” *Journal of Consumer Research*, 2022, *49* (2), 177–201.
- Kim, Donggwan, Zhenling Jiang, and Raphael Thomadsen**, “TV Advertising Effectiveness with Racial Minority Representation: Evidence from the Mortgage Market,” *Available at SSRN 4521178*, 2023.
- Kirgios, Erika L, Aneesh Rai, Edward H Chang, and Katherine L Milkman**, “When seeking help, women and racial/ethnic minorities benefit from explicitly stating their identity,” *Nature Human Behaviour*, 2022, *6* (3), 383–391.
- Krijnen, Tonny and Sofie Van Bauwel**, *Gender and media: Representing, producing, consuming*, Routledge, 2021.
- Overgoor, Gijs, Gokhan Yildirim, Yakov Bart, and Koen Pauwels**, “From Representation to Reception: Evaluating the Impact of Diversity in TV Advertising on Consumer Purchase Intention,” *Available at SSRN 4471248*, 2023.
- Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay**, “Scikit-learn: Machine Learning in Python,” *Journal of Machine Learning Research*, 2011, *12*, 2825–2830.
- UNICEF and UN Women**, “Dismantling gender stereotypes in media and advertising,” 2023 2023.
- Wang, Yang, Marco Shaojun Qin, Xueming Luo, and Yu Kou**, “Frontiers: How support for Black Lives Matter impacts consumer responses on social media,” *Marketing Science*, 2022, *41* (6), 1029–1044.

## A Gender Associations Survey

The survey study was voluntarily taken by students at a large university in Asia in exchange for course credits. The survey was open to 705 students and ran between February and April 2024. In total, 693 students completed the survey.

### A.1 Survey Outline

Study title: Advertising associations

**IRB/INFORMED CONSENT PAGE** Introductory text with IRB disclosure and link to PIS-C form.

**EXPLANATION PAGE** In this study, we are interested in your opinions about the content of some TV advertisements (“TV ads”).

In the next screens, we will show you a short text description of the content of a TV video advertisement. The text description will resemble a string of words or concepts separated by a slash symbol (/), describing advertising scenes in sequence. This is an example of such descriptions:

BOTTLE OF PERFUME/WATCH/LOGO

So for instance, this TV ad featured a bottle of perfume in the first scene, followed by a scene containing a watch, and closed with the scene of a brand logo.

**NEXT SCREEN – EXAMPLE AND PRACTICE ROUND** Next to each advertisement description, we will ask your opinion on what type of people are most likely to be associated with the content of the TV ad.

In particular, whether you think the ad content is most likely associated to a man (or group of men), a woman (or group of women), or both are equally likely to be associated with the ad content.

We ask you to express your opinion on a slider scale ranging from -1 (most likely associated to man/men) to 1 (most likely associated to woman/women). The mid-range point at 0 indicates that both groups are equally likely to be associated with the content of the ad.

Finally, you have an option to select “Not Applicable” if you think none of the two groups is likely to be associated to the content of the ad. Please try to rate the example description:

[Participants are shown 3 practice descriptions:

PEOPLE DANCING/CELL PHONE/TEXT/LOGO  
PERSON WITH CHAIN SAW/WORK SITE/LOGS  
IRONING DRESS/HANDS/5 IRONS/LOGO]

[Participants are shown a slider ranging from -1 to +1, like below:]



**MEASUREMENT QUESTIONS** Participants are shown 30 randomly selected TV advertising descriptions, from which we previously removed clearly gendered terms, and they are asked to indicate the most likely association using the same slider instrument as above.

## SOCIODEMOGRAPHICS

1. What is your age? [Under 18, 18-24, 25-34, 35-54, 55 or older]
2. What gender do you identify with? [Male, Female, Non-binary, Prefer not to say]
3. How often do you watch TV? [Never, Once a month or less, Once every 2 weeks, Once a week, Multiple times a week, Daily]
4. (Measurement of stereotypical beliefs, from Bordalo et al. 2019) Which gender, on average, knows more about each category of knowledge in general? [Verbal skills, cooking, business, rock and roll, the Kardashians, Disney movies, videogames, mathematics, emotion recognition, sports and games, arts and literature, cars]

**FEEDBACK** Do you have any feedback on the survey?

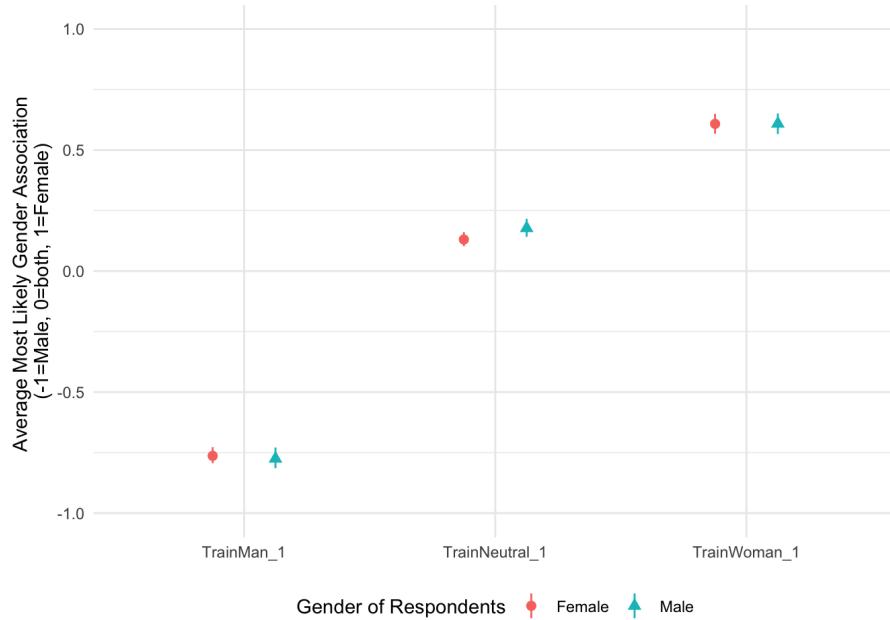
## A.2 Survey Descriptive Statistics

Of 693 students who completed the survey, 50.36% self-identified as male, 48.05% as female, and 1.59% preferred not to disclose their gender identity. 98.12% of the respondents was aged 18 – 24, 1.59% 25 – 34, and 0.29% was under 18 years old.

3.46% of the respondents reported that they watch TV daily. 12.55% watch TV multiple times a week, 15.01% once a week, 11.11% once every two weeks, 39.54% once a month or less, and 18.33% report never watching TV.

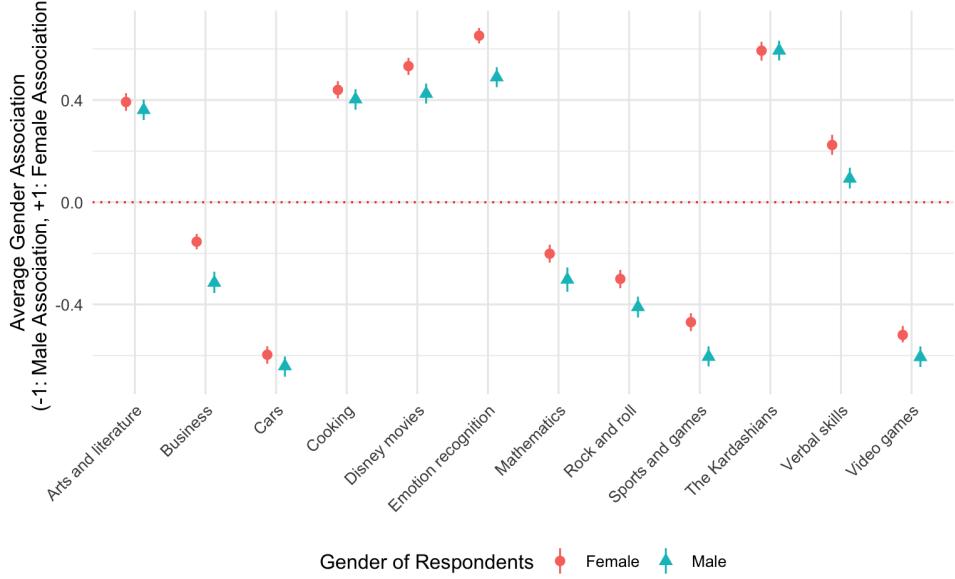
As illustrated above, participants were asked to score 3 ad creative descriptions as “training” for their task. These descriptions were chosen by the researchers as likely to elicit commonplace beliefs about gender associations (one neutral description, one more likely to be associated with men, and one more likely to be associated with women). Figure 13 shows that survey participants rated these training descriptions in line with the expected gender associations and that their beliefs did not significantly vary by gender of the respondents.

**Figure 13:** Association Scores for “Training” Ad Creative Descriptions – Gender Breakdown.



Respondents also answered questions about their own beliefs about the ability of people of different genders in different knowledge domains, based on Bordalo et al. (2019).

**Figure 14:** Survey Respondents' Average Scores for Gender Knowledge in Specific Categories.



## B Measuring Gender Role Model and Gender Context from TV Ad Creative Descriptions

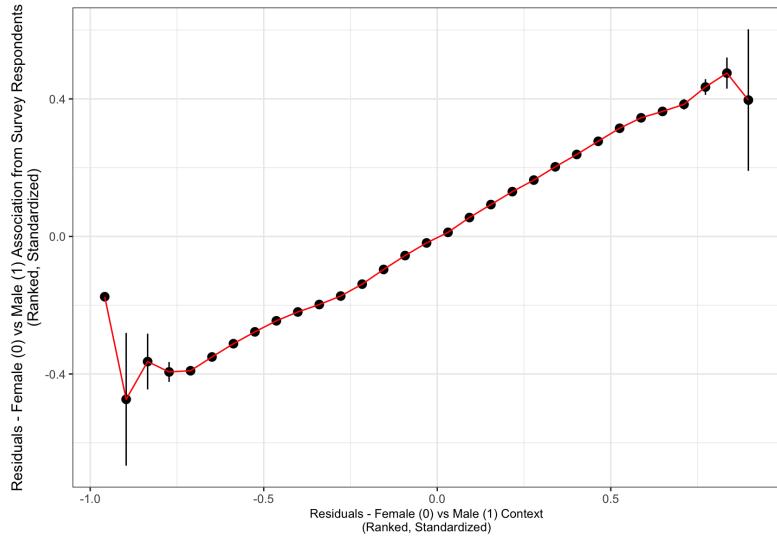
### B.1 List of Keywords

**Table 2:** Keywords used to compute gender role model indicator.

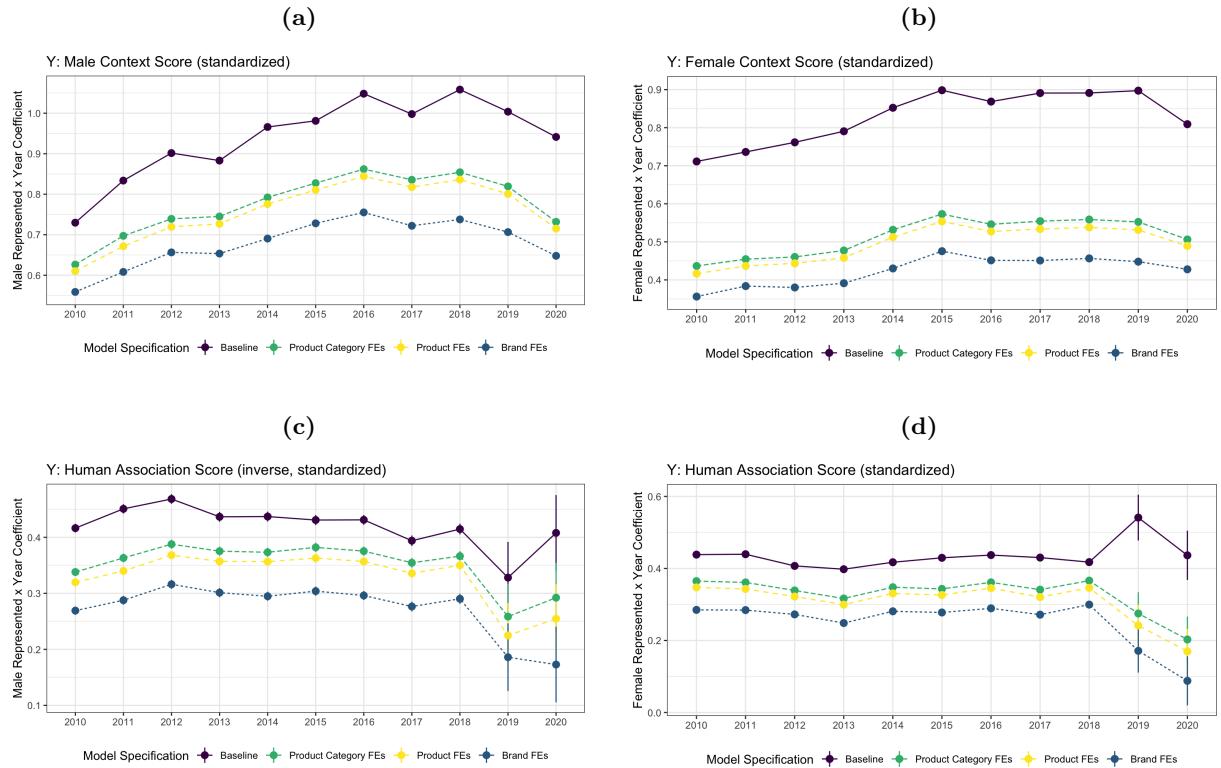
Male Keywords	'Man', 'Men', 'Boy', 'Boyfriend', 'Mans', 'Policemen', 'Sportsmen', 'Sportsman', 'Firemen', 'Fireman', 'Male', 'Father', 'Dad', 'Son', 'Actor', 'Husband', 'Waiter', 'Steward', 'Prince', 'Groom'
Female Keywords	'Girl', 'Woman', 'Lady', 'Girlfriend', 'Women', 'Womans', 'Ladies', 'Policewives', 'Wife', 'Sportswomen', 'Sportswoman', 'Firewomen', 'Firewoman', 'Female', 'Mother', 'Actress', 'Waitress', 'Stewardess', 'Daughter', 'Princess', 'Bride'

## C Additional Tables and Figures

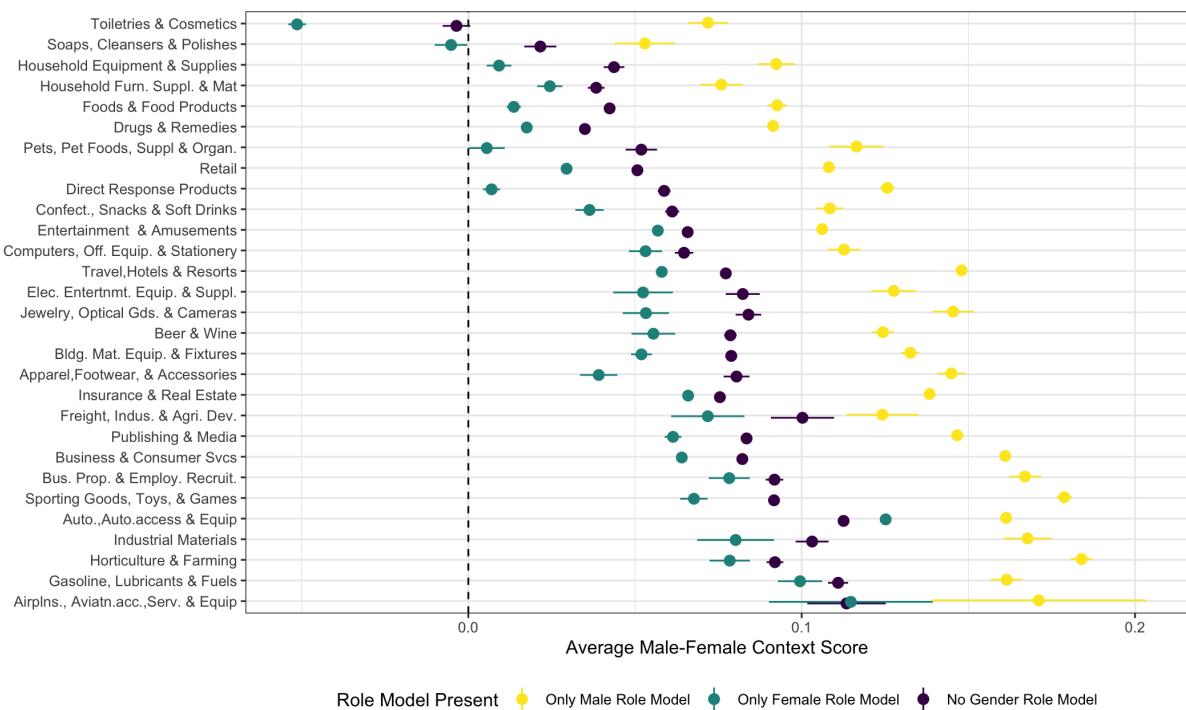
**Figure 15:** Residual Variation in Average Human Association Scores versus Residual Variation in Average Context Scores, Ranked and Scaled. The residuals are obtained from OLS regressions of the form  $\widehat{H}(a) = \alpha + \beta_b + \gamma_p + \varepsilon_{H(a)}$  and  $M(a) = \alpha + \beta_b + \gamma_p + \varepsilon_{M(a)}$ , where  $\beta_b, \gamma_p$  are brand and product fixed effects.



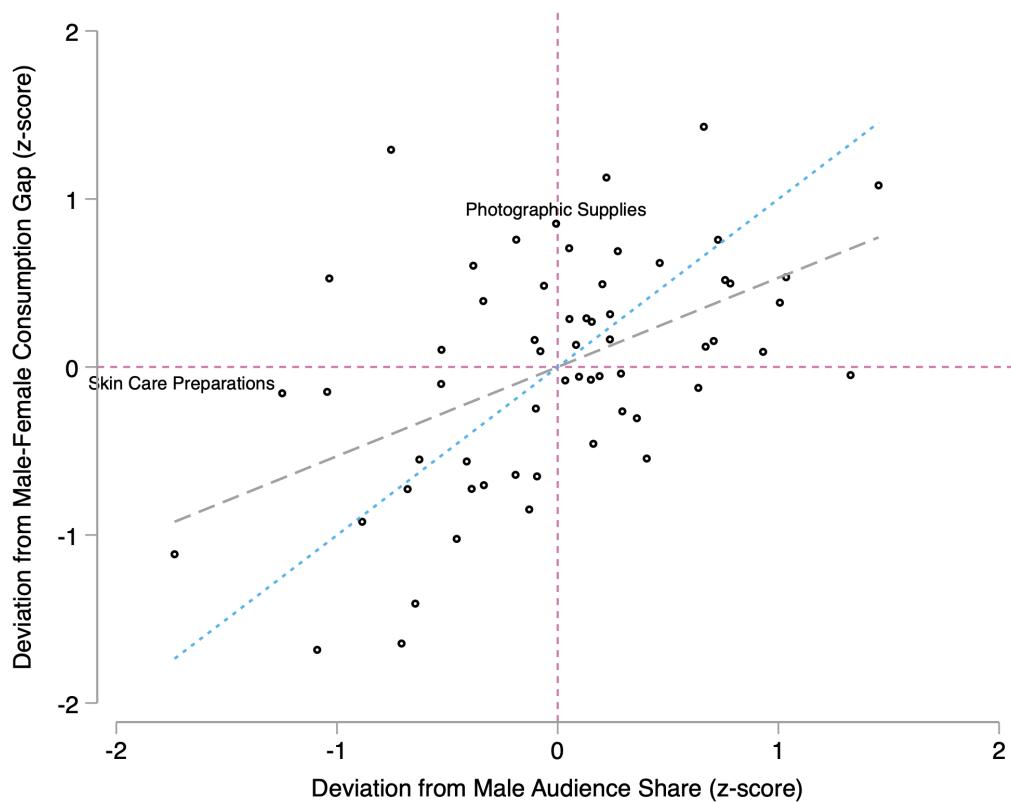
**Figure 16:** Yearly Representation of Role Models and Context Associations across TV Ads (2010-2020). Colored lines indicate linear models with fixed effects for brands (blue), product categories (green), and products (yellow). Regression specifications for (16a) and (16b) are of the form  $\tilde{M}(a) = \alpha_{\text{year}} \times \text{Male}(a) + \varepsilon$  and  $\tilde{M}(a) = \alpha_{\text{year}} \times \text{Female}(a) + \varepsilon | \text{Fixed Effects}$ , where  $\tilde{M}(a)$  is the standardized and centered data-driven context association score. Regression specifications for (16c) and (16d) are of the form  $\tilde{H}(a) = \alpha_{\text{year}} \times \text{Male}(a) + \varepsilon$  and  $\tilde{H}(a) = \alpha_{\text{year}} \times \text{Female}(a) + \varepsilon | \text{Fixed Effects}$ , where  $\tilde{H}(a)$  is the standardized and centered survey-based context association score.



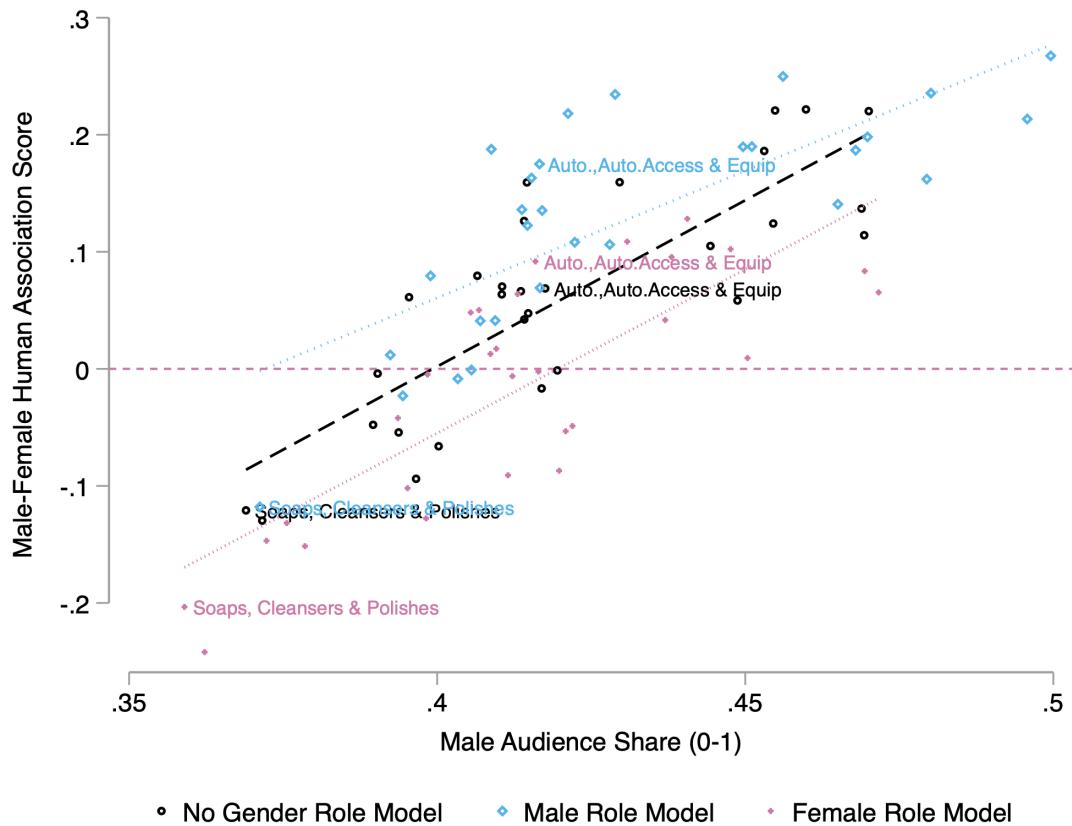
**Figure 17:** Yearly Representation of Role Models and (data-driven) Context Association across TV Ads and Product Industries. Colors indicate linear models estimated on ads with male role models (yellow), female role models (green), and neither gender as role model (purple).



**Figure 18:** Correlation between deviations of male–female representation from i) consumer base and ii) audience share ( $r=0.498$ ,  $p\text{-value}<0.001$ ,  $N=60$ ). Deviation is measured as the residual unexplained variation in male–female representation (as role models, and in context) after controlling for the male–female consumption gap, or the male audience share. Dashed grey line: linear fit. Dashed blue line:  $45^\circ$  line.



**Figure 19:** Average Male–Female Human Association Scores versus Average Male Audience Share by Industries (N=29) for Ads with Male (r=0.763), Female (r=0.796) or No Gender Roles (r=0.821).



**Figure 20:** Yearly Male–Female Human Association Scores of Major Diapers Brands (2010–2020).

